

(19)



**Евразийское
патентное
ведомство**

(11) **045713**

(13) **B1**

(12) ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ЕВРАЗИЙСКОМУ ПАТЕНТУ

(45) Дата публикации и выдачи патента
2023.12.20

(21) Номер заявки
201791558

(22) Дата подачи заявки
2016.02.10

(51) Int. Cl. **H04N 21/262** (2011.01)
H04N 21/438 (2011.01)
H04N 21/845 (2011.01)
H04N 21/854 (2011.01)
H04N 21/2343 (2011.01)
H04N 21/266 (2011.01)
H04N 21/44 (2011.01)

(54) СПОСОБ И КЛИЕНТСКОЕ УСТРОЙСТВО ДЛЯ ИЗВЛЕЧЕНИЯ МУЛЬТИМЕДИЙНЫХ ДАННЫХ ИЗ СЕРВЕРНОГО УСТРОЙСТВА

(31) 62/114,423; 62/183,054; 15/019,804

(32) 2015.02.10; 2015.06.22; 2016.02.09

(33) US

(43) 2017.12.29

(86) PCT/US2016/017325

(87) WO 2016/130657 2016.08.18

(71)(73) Заявитель и патентовладелец:
КВЭЛКОММ ИНКОРПОРЕЙТЕД
(US)

(72) Изобретатель:
Штокхаммер Томас, Ван Е-Куй (US)

(74) Представитель:
Медведев В.Н. (RU)

(56) ANONYMOUS: "Text of ISO/IEC CD 23001-6 Dynamic Adaptive Streaming over HTTP", 94. MPEG MEETING; 11-10-2010 - 15-10-2010; GUANGZHOU; (MOTION PICTURE EXPERT GROUP OR ISO/IEC JTC1/SC29/WG11), no. N11578, 21 October 2010 (2010-10-21), XP030018073, ISSN: 0000-0009, page 3-page 4, page 11, page 32-page 38, page 43-page 44

ZHUIE ZHAO ET AL. "Response for EE#5: Indexing Information and its Carriage for MPEG-2 TS", 95. MPEG MEETING; 24-1-2011 - 28-1-2011; DAEGU; (MOTION PICTURE EXPERT GROUP OR ISO/IEC JTC1/SC29/WG11), no. m18734, 25

November 2010 (2010-11-25), XP030047304, the whole document

GERARD FERNANDO ET AL. "Data format and delivery of side information for DASH streaming of MPEG-2 TS media", 95. MPEG MEETING; 24-1-2011 - 28-1-2011; DAEGU; (MOTION PICTURE EXPERT GROUP OR ISO/IEC JTC1/SC29/WG11), no. m18733, 25 November 2010 (2010-11-25), XP030047303, the whole document

YE-KUI WANG ET AL. "EE#6: Trick Mode and Random Access Signaling (TRA)", 94. MPEG MEETING; 11-10-2010 - 15-10-2010; GUANGZHOU; (MOTION PICTURE EXPERT GROUP OR ISO/IEC JTC1/SC29/WG11), no. M18016, 11 October 2010 (2010-10-11), XP030046606, page 11-page 20

US-A1-2013091251

US-A1-2011317760

US-A1-2012016965

ALEX GILADI: "Integrated text for 23009-3", 108. MPEG MEETING; 31-3-2014 - 4-4-2014; VALENCIA; (MOTION PICTURE EXPERT GROUP OR ISO/IEC JTC1/SC29/WG11), no. m33470, 30 March 2014 (2014-03-30), XP030061921, page 2-page 10, page 43-page 45

THOMAS STOCKHAMMER: "SISSI Solution Proposal", 112. MPEG MEETING; 22-6-2015 - 26-6-2015; WARSAW; (MOTION PICTURE EXPERT GROUP OR ISO/IEC JTC1/SC29/WG11), no. m36598, 23 June 2015 (2015-06-23), XP030064966, the whole document

(57) Клиентское устройство включает в себя один или более процессоров, выполненных с возможностью определять, из файла манифеста, множество типов сегментов, включенных в представление мультимедийного контента, причем одна или более функций обеспечиваются посредством каждого из типов сегментов и позиций сегментов, соответствующих каждому из типов сегментов в представлении, при этом, по меньшей мере, один из типов сегментов обеспечивает точку, в которой можно начинать извлечение данных из представления, определять, из файла манифеста, сегмент представления, соответствующий типу, который обеспечивает точку, в которой можно начинать извлечение данных из представления, и извлекать определенный сегмент из представления.

B1

045713

045713 B1

Данная заявка притязает на приоритет предварительной заявки на патент (США) номер 62/114423, поданной 10 февраля 2015 года, и предварительной заявки на патент (США) номер 62/183054, поданной 22 июня 2015 года, содержимое каждой из которых полностью содержится в данном документе по ссылке.

Область техники, к которой относится изобретение

Изобретение относится к хранению и транспортировке кодированных видеоданных.

Уровень техники

Возможности цифрового видео могут быть встроены в широкий диапазон устройств, включающих в себя цифровые телевизоры, системы цифровой прямой широковещательной передачи, беспроводные широковещательные системы, персональные цифровые устройства (PDA), переносные или настольные компьютеры, цифровые камеры, цифровые записывающие устройства, цифровые мультимедийные проигрыватели, устройства для видеоигр, консоли для видеоигр, сотовые или спутниковые радиотелефоны, устройства видеоконференц-связи и т.п. Цифровые видеоприемники реализуют такие технологии сжатия видео, как технологии, описанные в стандартах, заданных посредством MPEG-2, MPEG-4, ITU-T H.263, ITU-T H.264/MPEG-4, часть 10, усовершенствованное кодирование видео (AVC), ITU-T H.265/стандарта высокоэффективного кодирования видео (HEVC) и расширений таких стандартов, для того чтобы более эффективно передавать и принимать цифровую видеосообщения.

Технологии сжатия видео выполняют пространственное прогнозирование и/или временное прогнозирование для того, чтобы уменьшать или удалять избыточность, внутренне присутствующую в видеопоследовательностях. Для кодирования видео на основе блоков, видеокادر или серия последовательных макроблоков может быть сегментирована на макроблоки. Каждый макроблок может быть дополнительно сегментирован. Макроблоки во внутренне кодированном (I-) кадре или серии внутренне кодированных (I-) последовательных макроблоков кодируются с использованием пространственного прогнозирования относительно соседних макроблоков. Макроблоки во взаимно кодированном (P- или B-) кадре или серии взаимно кодированных (P- или B-) последовательных макроблоков могут использовать пространственное прогнозирование относительно соседних макроблоков в идентичном кадре или серии последовательных макроблоков либо временное прогнозирование относительно других опорных кадров.

После того, как видеоданные кодированы, видеоданные могут быть пакетированы для передачи или хранения. Видеоданные могут асемблироваться в видеофайл, соответствующий любому из множества стандартов, таких как базовый формат мультимедийных файлов Международной организации по стандартизации (ISO) и его расширения, к примеру, AVC.

Сущность изобретения

В общем, это раскрытие сущности описывает технологии, которые могут использоваться для того, чтобы достигать потоковой передачи видео (и/или других мультимедийных данных) с низкой задержкой. Например, мультимедийный контент может включать в себя множество представлений, которые выступают в качестве альтернатив друг другу. В соответствии с технологиями этого раскрытия сущности, одно представление может включать в себя относительно частые точки доступа к потоку (SAP), тогда как другое альтернативное представление может включать в себя относительно нечастые SAP. Файл манифеста (такой как описание мультимедийного представления (MPD) динамической адаптивной потоковой передачи по HTTP (DASH)) может передавать в сигналах типы сегментов (или форматы, которым соответствуют сегменты), а также местоположения таких сегментов (или относительные частоты, с которыми такие сегменты возникают в соответствующем представлении. Клиентское устройство может использовать файл манифеста для того, чтобы определять одно из представлений, имеющих относительно частые SAP, и затем извлекать сегменты или части сегментов из этого представления до тех пор, пока SAP не станет доступна из другого целевого представления. Целевое представление может иметь относительно более высокое качество вследствие наличия меньшего числа (т.е. менее частых) SAP. В некоторых примерах, различные представления могут быть доступными через различные механизмы извлечения, такие как одноадресная передача или широковещательная передача. Например, начальное представление может быть доступным через одноадресную передачу, тогда как целевое представление может быть доступным через широковещательную передачу.

В одном примере, способ включает в себя определение, из файла манифеста, множества типов сегментов, включенных в представление мультимедийного контента, причем одна или более функций обеспечиваются посредством каждого из типов сегментов и позиций сегментов, соответствующих каждому из типов сегментов в представлении, при этом, по меньшей мере, один из типов сегментов обеспечивает точку, в которой можно начинать извлечение данных из представления, определение, из файла манифеста, сегмента представления, соответствующего типу, который обеспечивает точку, в которой можно начинать извлечение данных из представления, и извлечение определенного сегмента из представления.

В другом примере, клиентское устройство для извлечения мультимедийных данных включает в себя один или более процессоров, выполненных с возможностью определять, из файла манифеста, множество типов сегментов, включенных в представление мультимедийного контента, причем одна или более функций обеспечиваются посредством каждого из типов сегментов и позиций сегментов, соответствующих каждому из типов сегментов в представлении, при этом, по меньшей мере, один из типов сегментов

рой можно начинать извлечение данных из представления, отправлять файл манифеста в клиентское устройство и отправлять сегмент, который обеспечивает точку, в которой можно начинать извлечение данных из представления, в клиентское устройство в ответ на запрос из клиентского устройства на предмет сегмента, соответствующего типу, который обеспечивает точку, в которой можно начинать извлечение данных из представления.

Подробности одного или более примеров изложены на прилагаемых чертежах и в нижеприведенном описании. Другие признаки, цели и преимущества должны становиться очевидными из описания и чертежей и из формулы изобретения.

Краткое описание чертежей

Фиг. 1 является концептуальной схемой, иллюстрирующей примерный вариант использования для быстрого присоединения к потоку.

Фиг. 2 является диаграммой Венна, иллюстрирующей взаимосвязи между различными типами мультимедийных сегментов.

Фиг. 3 является концептуальной схемой, иллюстрирующей примерную структуру представления и базового формата мультимедийных файлов ISO (BMFF) файл.

Фиг. 4 является блок-схемой, иллюстрирующей примерную систему, которая реализует технологии для потоковой передачи мультимедийных данных по сети.

Фиг. 5А является концептуальной схемой, иллюстрирующей элементы примерного мультимедийного контента.

Фиг. 5В является концептуальной схемой, иллюстрирующей примерный контент описания мультимедийного представления в соответствии с технологиями этого раскрытия сущности.

Фиг. 6 является блок-схемой, иллюстрирующей элементы примерного видеофайла, которые могут соответствовать сегменту представления, такому как один из сегментов по фиг. 5А.

Фиг. 7 является концептуальной схемой, иллюстрирующей примерное предложение сегментов для варианта использования согласно технологиям этого раскрытия сущности.

Фиг. 8 является концептуальной схемой, иллюстрирующей вариант использования, включающий в себя быструю настройку с помощью масштабируемого HEVC (SHVC) в соответствии с технологиями этого раскрытия сущности.

Фиг. 9 является концептуальной схемой, иллюстрирующей примерный вариант использования, включающий в себя быструю настройку с помощью точки доступа к потоку (SAP) типа 3 в соответствии с технологиями этого раскрытия сущности.

Фиг. 10 является концептуальной схемой, иллюстрирующей примерный вариант использования, включающий в себя быструю настройку и гибридизацию.

Фиг. 11 является концептуальной схемой, иллюстрирующей примерный вариант использования, включающий в себя быструю настройку, гибридизацию и открытые GOP.

Фиг. 12 является концептуальной схемой, иллюстрирующей другой примерный вариант использования, включающий в себя быструю настройку и гибридизацию с открытыми GOP.

Фиг. 13 является концептуальной схемой, иллюстрирующей примерный вариант использования, включающий в себя быструю настройку и очень низкую задержку.

Фиг. 14 является концептуальной схемой, иллюстрирующей другой примерный вариант использования, включающий в себя быструю настройку и очень низкую задержку.

Фиг. 15 является блок-схемой последовательности операций, иллюстрирующей примерный способ для извлечения сегмента представления мультимедийного контента в соответствии с технологиями этого раскрытия сущности.

Подробное описание изобретения

В общем, это раскрытие сущности описывает технологии для потоковой передачи видео с низкой задержкой, например, на основе мультимедийного контента, отформатированного согласно базовому формату мультимедийных файлов ISO (ISO/BMFF) и динамической адаптивной потоковой передаче по HTTP (DASH). DASH описывается, например, в документе 3rd Generation Partnership Project; Technical Specification Group Services and System Aspects; Transparent end-to-end packet switched streaming service (PSS); 3GPP file format (3GP) (Release 12) V12.2.0, декабрь 2013 года. Это раскрытие сущности описывает различные способы для задания и передачи в сигналах данных, которые могут соответствовать новому DASH-профилю (например, усовершенствованному профилю передач вживую), и некоторых новых типов мультимедийных сегментов, которые могут обеспечивать потоковую передачу видео с низкой задержкой, включающую в себя уменьшенные времена получения каналов и переключения каналов при широкоэмитальной передаче и многоадресной передаче, при потенциальном одновременном обеспечении структур по стандарту высокоэффективного кодирования видео.

Стандарты кодирования видео включают в себя ITU-T H.261, ISO/IEC MPEG-1 Visual, ITU-T H.262 или ISO/IEC MPEG-2 Visual, ITU-T H.263, ISO/IEC MPEG-4 Visual, ITU-T H.264 или ISO/IEC MPEG-4 AVC, в том числе его расширения масштабируемого кодирования видео (SVC) и кодирования многовидеового видео (MVC), и стандарт высокоэффективного кодирования видео (HEVC), также известный как ITU-T H.265 и ISO/IEC 23008-2, в том числе его расширение масштабируемого кодирования (т.е. стан-

дарт масштабируемого высокоэффективного кодирования видео, SHVC) и многовидовое расширение (т.е. стандарт многовидового высокоэффективного кодирования видео, MV-HEVC).

Произвольный доступ означает декодирование потока видеобитов, начиная с кодированного изображения, которое не представляет собой первое кодированное изображение в потоке битов. Произвольный доступ к потоку битов может использоваться во многих видеоприложениях, таких как широкоэвещательная передача и потоковая передача, например, для настройки пользователей на программу в любое время, переключения между различными каналами, перехода в конкретные части видео или переключения на другой поток битов для адаптации потоков (скорости передачи битов, частоты кадров, пространственного разрешения и т.д.). Этот признак может обеспечиваться посредством вставки изображений с произвольным доступом или точек произвольного доступа, много раз с регулярными интервалами, в поток видеобитов.

Сращивание потоков битов означает конкатенацию двух или более потоков битов или их частей. Например, к первому потоку битов может добавляться второй поток битов, возможно с некоторыми модификациями в одном или обоих из потоков битов, с тем чтобы формировать сращиваемый поток битов. Первое кодированное изображение во втором потоке битов также упоминается как точка сращивания. Следовательно, изображения после точки сращивания в сращиваемом потоке битов исходят из второго потока битов, тогда как изображения, предшествующие точке сращивания в сращиваемом потоке битов, исходят из первого потока битов.

Сращивание потоков битов может выполняться посредством модулей сращивания потоков битов. Модули сращивания потоков битов зачастую являются упрощенными и гораздо менее интеллектуальными, чем кодеры. Например, модули сращивания потоков битов могут не быть оснащены возможностями энтропийного декодирования и кодирования.

Переключение потоков битов может использоваться в окружениях адаптивной потоковой передачи. Операция переключения потоков битов, выполняемая в определенном изображении в потоке битов, на который выполняется переключение, фактически представляет собой операцию сращивания потоков битов, в которой точка сращивания представляет собой точку переключения потоков битов, т.е. первое изображение из потока битов, на который выполняется переключение. Отдельные представления также могут упоминаться как (или обеспечивать) соответствующие потоки битов.

Изображения на основе мгновенного обновления при декодировании (IDR), как указано в стандарте ITU-T H.264/AVC (усовершенствованного кодирования видео) или в стандарте высокоэффективного кодирования видео (HEVC), могут использоваться для произвольного доступа. Тем не менее, поскольку изображения, идущие после IDR-изображения в порядке декодирования, не могут использовать изображения, декодированные до IDR-изображения, для обращения (для межкадрового прогнозирования), потоки битов, основывающиеся на IDR-изображениях для произвольного доступа, могут иметь значительно меньшую эффективность кодирования.

Чтобы повысить эффективность кодирования, принцип изображений на основе чистого произвольного доступа (CRA) введен в HEVC, чтобы обеспечивать возможность изображениям, которые идут после CRA-изображения в порядке декодирования, но предшествуют ему в порядке вывода, использовать изображения, декодированные перед CRA-изображением, для обращения. Изображения, которые идут после CRA-изображения в порядке декодирования, но предшествуют CRA-изображению в порядке вывода, упоминаются как начальные изображения, ассоциированные с CRA-изображением (или начальные изображения по отношению к CRA-изображению). Начальные изображения по отношению к CRA-изображению являются корректно декодируемыми, если декодирование начинается с IDR- или CRA-изображения перед текущим CRA-изображением. Тем не менее, начальные изображения по отношению к CRA-изображению могут быть недекодируемыми, когда возникает произвольный доступ из CRA-изображения. Следовательно, начальные изображения типично отбрасываются во время декодирования с произвольным доступом. Чтобы предотвращать распространение ошибки из опорных изображений, которые могут быть недоступными в зависимости от того, где начинается декодирование, все изображения, которые идут после CRA-изображения как в порядке декодирования, так и порядке вывода, не должны использовать изображения, которые предшествуют CRA-изображению либо в порядке декодирования, либо в порядке вывода (которые включают в себя начальные изображения), для обращения.

Принцип изображения на основе доступа к нерабочей ссылке (BLA) дополнительно введен в HEVC после введения CRA-изображений и основан на принципе CRA-изображений. BLA-изображение типично исходит из сращивания потоков битов в позиции CRA-изображения, и в сращиваемом потоке битов CRA-изображение точки сращивания изменяется на BLA-изображение.

IDR-изображения, CRA-изображения и BLA-изображения совместно упоминаются в качестве изображений на основе точки произвольного доступа (RAP). IDR-изображения соответствуют так называемым RAP на основе закрытой группы изображений (GOP), тогда как CRA- и BLA-изображения соответствуют традиционно так называемым RAP на основе открытой группы изображений (GOP).

Одно отличие между BLA-изображениями и CRA-изображениями заключается в следующем. Для CRA-изображения, ассоциированные начальные изображения являются корректно декодируемыми, если декодирование начинается с RAP-изображения перед CRA-изображением в порядке декодирования, и

могут быть некорректно декодируемыми, когда возникает произвольный доступ из CRA-изображения (т.е. когда декодирование начинается с CRA-изображения, или другими словами, когда CRA-изображение представляет собой первое изображение в потоке битов). Для BLA-изображения, ассоциированные начальные изображения могут быть недекодируемыми во всех случаях, даже когда декодирование начинается с RAP-изображения перед BLA-изображением в порядке декодирования.

Стандарты форматов файлов включают в себя базовый формат мультимедийных файлов ISO (ISO-BMFF, ISO/IEC 14496-12) и другие, извлеченные из ISO-BMFF, в том числе MPEG-4-формат файлов (ISO/IEC 14496-14), 3GPP-формат файлов (3GPP TS 26.244) и AVC-формат файлов (ISO/IEC 14496-15).

ISO-BMFF используется в качестве базиса для многих форматов инкапсуляции кодека, таких как AVC-формат файлов, а также для многих форматов мультимедийных контейнеров, таких как MPEG-4-формат файлов, 3GPP-формат файлов (3GP) и DVB-формат файлов.

В дополнение к непрерывному мультимедиа, такому как аудио и видео, статическое мультимедиа, такое как изображения, а также метаданные могут сохраняться в файле, соответствующем ISO-BMFF. Файлы, структурированные согласно ISO-BMFF, могут использоваться для многих целей, включающих в себя локальное воспроизведение мультимедийных файлов, прогрессивную загрузку удаленного файла, сегменты для динамической адаптивной потоковой передачи по HTTP (DASH), контейнеры для контента, который должен передаваться в потоковом режиме, и инструкции по его пакетированию, и запись принимаемых мультимедийных потоков в реальном времени.

Поле представляет собой элементарную синтаксическую структуру в ISO-BMFF, включающую в себя четырехсимвольный кодированный тип поля, число байтов поля и рабочие данные. ISO-BMFF-файл состоит из последовательности полей, и поля могут содержать другие поля. Кинополе (moov) содержит метаданные для непрерывных мультимедийных потоков, присутствующих в файле, при этом каждый из них представляется в файле в качестве дорожки. Метаданные для дорожки включаются в поле дорожек (trak), тогда как мультимедийный контент дорожки включается либо в поле мультимедийных данных (mdat), либо непосредственно в отдельный файл. Мультимедийный контент для дорожек состоит из последовательности выборок, таких как единицы аудио- или видеодоступа.

ISO-BMFF указывает следующие типы дорожек: мультимедийная дорожка, которая содержит элементарный мультимедийный поток, дорожка с подсказками, которая или включает в себя инструкции по передаче мультимедиа или представляет поток принимаемых пакетов, и дорожка синхронизированных по времени метаданных, которая содержит метаданные с временной синхронизацией.

Хотя первоначально спроектирован для хранения, ISO-BMFF оказался очень ценным для потоковой передачи, например, для прогрессивной загрузки или DASH. Для целей потоковой передачи, могут использоваться кинофрагменты, заданные в ISO-BMFF.

Метаданные для каждой дорожки включают в себя список записей описания выборок, каждая из которых обеспечивает формат кодирования или инкапсуляции, используемый в дорожке, и данные инициализации, необходимые для обработки этого формата. Каждая выборка ассоциирована с одной из записей описания выборок дорожки.

ISO-BMFF обеспечивает указание конкретных для выборки метаданных с помощью различных механизмов. Конкретные поля в поле таблиц выборок (stbl) стандартизированы, с тем чтобы реагировать на общие потребности. Например, поле выборок синхроимпульсов (stss) используется для того, чтобы перечислять выборки с произвольным доступом дорожки. Механизм группировки выборок обеспечивает преобразование выборок согласно четырехсимвольному типу группировки в группы выборок, совместно использующих идентичное свойство, указываемое в качестве записи описания групп выборок в файле. Несколько типов группировки указаны в ISO-BMFF.

Технологии этого раскрытия сущности могут применяться к видеофайлам, соответствующим видеоданным, инкапсулированным согласно любому из ISO-BMFF, формата файлов по стандарту масштабируемого кодирования видео (SVC), формата файлов по стандарту усовершенствованного кодирования видео (AVC), формата файлов по стандарту Партнерского проекта третьего поколения (3GPP) и/или формата файлов по стандарту кодирования многовидового видео (MVC) либо других аналогичных форматов видеофайлов.

ISO/IEC 23001-7 задает общее шифрование для базового формата мультимедийных файлов ISO. В случае этого стандарта, шифрование основано на элементарном потоке. Помимо этого, стандарт обеспечивает возможность AES-128 CTR- и CBC-режима. Чтобы дешифровать мультимедиа в точке произвольного доступа, требуется вся связанная с DRM информация, включающая в себя конкретную для схемы защиты информацию, а также векторы инициализации.

Динамическая адаптивная потоковая передача по HTTP (DASH), указываемая в ISO/IEC 23009-1, представляет собой стандарт для приложений (адаптивной) потоковой HTTP-передачи данных. Она главным образом указывает формат описания мультимедийного представления (MPD), также, в общем, называемого в качестве файла манифеста и формата мультимедийных сегментов. MPD описывает мультимедиа, доступное на сервере, и позволяет DASH-клиенту автономно загружать версию мультимедиа в то время в мультимедиа, в которое ему нужно.

Примерная процедура для потоковой HTTP-передачи на основе DASH включает в себя следующие

этапы:

1) Клиент получает MPD потокового контента, например, фильма. MPD включает в себя информацию относительно различных альтернативных представлений, например, скорость передачи битов, разрешение видео, частоту кадров, аудиоязык, для потокового контента, а также URL-адреса HTTP-ресурсов (сегмента инициализации и мультимедийных сегментов).

2) На основе информации в MPD и локальной информации клиента, например, полосы пропускания сети, характеристик декодирования/отображения и настроек пользователя, клиент запрашивает требуемое представление(я), один сегмент (или часть его) за раз.

3) Когда клиент обнаруживает изменение полосы пропускания сети, клиент запрашивает сегменты другого представления с лучше совпадающей скоростью передачи битов, в идеале начиная с сегмента, который начинается с точки произвольного доступа.

Во время "сеанса" потоковой HTTP-передачи, чтобы отвечать на пользовательский запрос, с тем чтобы выполнять поиск назад до прошлой позиции или вперед до будущей позиции, клиент запрашивает прошлые или будущие сегменты, начиная с сегмента, который находится близко к требуемой позиции и который в идеале начинается с точки произвольного доступа. Пользователь также может запрашивать ускоренно перематывать вперед контент, что может быть реализовано посредством запрашивания данных в достаточной степени для декодирования только внутренне кодированных видеозображений или только временного поднабора видеопотока.

Последняя спецификация ISO/BMFF указывает шесть типов точек доступа к потоку (SAP) для использования с DASH. Первые два SAP-типа (типы 1 и 2) соответствуют IDR-изображениям в H.264/AVC и HEVC. Третий SAP-тип (тип 3) соответствует точкам произвольного доступа в форме открытой GOP, а следовательно, BLA- или CRA-изображениям в HEVC.

В потоковой HTTP-передаче, например, согласно DASH, часто используемые операции включают в себя HEAD, GET и частичный GET. Операция HEAD извлекает заголовок файла, ассоциированного с данным универсальным указателем ресурса (URL-адресом) или универсальным именем ресурса (URN), без извлечения рабочих данных, ассоциированных с URL-адресом или URN. Операция GET извлекает целый файл, ассоциированный с данным URL-адресом или URN. Операция частичного GET принимает байтовый диапазон в качестве входного параметра и извлекает непрерывное число байтов файла, при этом число байтов соответствует принимаемому байтовому диапазону. Таким образом, кинофрагменты могут обеспечиваться для потоковой HTTP-передачи, поскольку операция частичного GET может получать один или более отдельных кинофрагментов. В кинофрагменте, может быть предусмотрено несколько фрагментов дорожек для различных дорожек. В потоковой HTTP-передаче, мультимедийное представление может представлять собой структурированную совокупность данных, которая является доступной для клиента. Клиент может запрашивать и загружать информацию в виде мультимедийных данных, чтобы представлять услугу потоковой передачи пользователю.

В примере потоковой передачи 3GPP-данных с использованием потоковой HTTP-передачи, может быть предусмотрено несколько представлений для видео- и/или аудиоданных мультимедийного контента. Как пояснено ниже, различные представления могут соответствовать различным характеристикам кодирования (например, различным профилям или уровни стандарта кодирования видео), различным стандартам кодирования или расширениям стандартов кодирования (к примеру, многовидовым и/или масштабируемым расширениям) или различным скоростям передачи битов. Манифест таких представлений может задаваться в структуре данных описания мультимедийного представления (MPD). Мультимедийное представление может соответствовать структурированной совокупности данных, которая является доступной для клиентского устройства потоковой HTTP-передачи. Клиентское устройство потоковой HTTP-передачи может запрашивать и загружать информацию в виде мультимедийных данных, чтобы представлять услугу потоковой передачи пользователю клиентского устройства. Мультимедийное представление может описываться в структуре MPD-данных, которая может включать в себя обновления MPD.

Мультимедийное представление может содержать последовательность одного или более периодов. Периоды могут задаваться посредством элемента Period в MPD. Каждый период может иметь атрибут start в MPD. MPD может включать в себя атрибут start и атрибут availableStartTime в течение каждого периода. Для услуг передач вживую, сумма атрибута start периода и MPD-атрибута availableStartTime может указывать время доступности периода в UTC-формате, в частности, первый мультимедийный сегмент каждого представления в соответствующий период. Для услуг по запросу, атрибут start первого периода может быть равным 0. Для любого другого периода, атрибут start может указывать сдвиг по времени между начальным временем соответствующего периода относительно начального времени первого периода. Каждый период может длиться до начала следующего периода или до конца мультимедийного представления в случае последнего периода. Начальные времена периода могут быть точными. Они могут отражать фактическую временную синхронизацию, получающуюся в результате воспроизведения мультимедиа всех предшествующих периодов.

Каждый период может содержать одно или более представлений для идентичного мультимедийного контента. Представление может быть одной из ряда альтернативных кодированных версий аудио- или

видеоданных. Представления могут отличаться посредством типов кодирования, например, посредством скорости передачи битов, разрешения и/или кодека для видеоданных и скорости передачи битов, языка и/или кодека для аудиоданных.

Термин "представление" может использоваться для того, чтобы означать секцию кодированных аудио- или видеоданных, соответствующих конкретному периоду мультимедийного контента и кодированных конкретным способом.

Представления конкретного периода могут назначаться группе, указываемой посредством атрибута в MPD, указывающем адаптивный набор, которому принадлежат представления.

Представления в идентичном адаптивном наборе, в общем, считаются альтернативами друг другу в том, что клиентское устройство может динамически и прозрачно переключаться между этими представлениями, например, чтобы выполнять адаптацию полосы пропускания. Например, каждое представление видеоданных для конкретного периода может назначаться идентичному адаптивному набору, так что любое из представлений может выбираться для декодирования, чтобы представлять мультимедийные данные, к примеру, видеоданные или аудиоданные, мультимедийного контента для соответствующего периода. Мультимедийный контент в одном периоде может быть представлен либо посредством одного представления из группы 0, если присутствует, либо посредством комбинации самое большее одного представления из каждой ненулевой группы, в некоторых примерах. Данные временной синхронизации для каждого представления периода могут выражаться относительно начального времени периода.

Представление может включать в себя один или более сегментов. Каждое представление может включать в себя сегмент инициализации, или каждый сегмент представления может самоинициализироваться. Когда присутствует, сегмент инициализации может содержать информацию инициализации для осуществления доступа к представлению. В общем, сегмент инициализации не содержит мультимедийные данные. К сегменту можно уникально обращаться посредством идентификатора, такого как универсальный указатель ресурса (URL-адрес), универсальное имя ресурса (URN) или универсальный идентификатор ресурса (URI). MPD может обеспечивать идентификаторы для каждого сегмента. В некоторых примерах, MPD также может обеспечивать байтовые диапазоны в форме атрибута range, которые могут соответствовать данным для сегмента в файле, доступном посредством URL-адреса, URN или URI.

Различные представления могут выбираться для практически одновременного извлечения для различных типов мультимедийных данных. Например, клиентское устройство может выбирать аудиопредставление, видеопредставление и синхронизированное по времени текстовое представление, из которых можно извлекать сегменты. В некоторых примерах, клиентское устройство может выбирать конкретные адаптивные наборы для выполнения адаптации полосы пропускания. Иными словами, клиентское устройство может выбирать адаптивный набор, включающий в себя видеопредставления, адаптивный набор, включающий в себя аудиопредставления, и/или адаптивный набор, включающий в себя синхронизированный по времени текст. Альтернативно, клиентское устройство может выбирать адаптивные наборы для определенных типов мультимедиа (например, видео) и непосредственно выбирать представления для других типов мультимедиа (например, аудио и/или синхронизированного по времени текста).

Различные проблемы могут возникать в традиционных DASH-технологиях. Например, для услуг потоковой передачи видео с низкой задержкой, таких как распространение услуги передач вживую с низкой задержкой, релевантно, что каждый сегмент может формироваться максимально быстро, чтобы становиться доступным на исходном сервере. Другими словами, короткие сегменты необходимы в таких сценариях. В настоящее время, предусмотрено два варианта для создания коротких сегментов:

1) Использовать ISOBMFF-профиль передач вживую: Это означает то, что каждый сегмент должен начинаться с SAP типа 1 или 2, но все сегменты должны иметь идентичную длительность в одном адаптивном наборе. Другими словами, IDR-изображения должны использоваться для того, чтобы обеспечивать RAP, RAP в форме открытой GOP, которые соответствуют SAP-типу 3, не могут использоваться. Следовательно, эффективность кодирования видео должна снижаться.

2) Использовать главный ISOBMFF-профиль: Тем не менее, это означает то, что передача сигналов на основе MPD по точкам переключения (SAP-тип 1 или 2) невозможна, и клиент должен синтаксически анализировать сегменты, чтобы узнавать то, как осуществлять доступ к выборке.

Помимо этого, может возникать проблема перегрузки сегментов. Иными словами, в базовой DASH-спецификации, сегменты представляют собой единицы доставки, которые должны включать в себя целое число кинофрагментов. Без потери общности допустим, что сегмент содержит один кинофрагмент. Сами кинофрагменты имеют ограничения только с точки зрения обеспечения целого числа выборок в порядке декодирования.

В базовой DASH, сегменты могут формироваться в целях создания адресуемых и доставляемых единиц без дополнительных ограничений. Тем не менее, в ограниченных профилях (например, в ISO-профиле передач вживую), сегменты одновременно используются для обеспечения возможности переключения представлений. Последнее добавляет значимые ограничения:

каждый сегмент должен начинаться с закрытой GOP;

сегменты не должны прекрываться во времени представления в пределах одного представления.

Эти два ограничения приводят к уменьшенной эффективности кодирования, в частности, если сег-

менты являются относительно короткими.

Кроме того, для широковещательных вариантов применения, произвольный доступ в единице доставки является релевантным. Длительность сегментов определяет время произвольного доступа, которое является релевантным для получения каналов и переключения каналов. Для произвольного доступа, более эффективная открытая GOP является достаточной, и сегменты могут даже иметь перекрытие по времени представления в некоторой степени, которое может приводить к уменьшенному качеству воспроизведения при доступе (некоторым отброшенным кадрам), но по-прежнему обеспечивать возможность быстрого доступа к потоку.

Технологии этого раскрытия сущности, как пояснено ниже, могут разрешать различные функциональные аспекты сегмента и различать сегменты на различные классы.

Фиг. 1 является концептуальной схемой, иллюстрирующей примерный вариант использования для быстрого присоединения к потоку. В этом примере, некоторые сегменты доступны через широковещательную передачу, тогда как другие сегменты доступны через одноадресную передачу. В частности, сегменты с пометкой "8" и "9" доступны через широковещательную передачу, тогда как сегменты с пометкой 7A 7D, 8A-8D и 9A-9D доступны через одноадресную передачу. В этом примере использования, клиентское устройство извлекает сегменты 7D и 8A-8D через одноадресную передачу (при этом сегменты 8A-8D включают в себя мультимедийные данные, идентичные мультимедийным данным сегмента 8, доступного через широковещательную передачу), а затем принимает сегмент 9 через широковещательную передачу. В частности, клиентское устройство настраивается на широковещательную передачу во время 2 настройки, т.е. в ходе передачи сегмента 8 через широковещательную передачу. Следовательно, клиентское устройство не может принимать сегмент 8 через широковещательную передачу, так что вместо этого, клиентское устройство извлекает сегменты 7D и 8A-8D перед приемом сегмента 9 через широковещательную передачу. Таким образом, клиентское устройство переключается с широковещательной передачи на одноадресную передачу после извлечения сегмента 8D. Соответственно, при воспроизведении мультимедийных данных, клиентское устройство воспроизводит мультимедийные данные из сегментов 7D и 8A-8D (принимаемых через одноадресную передачу), затем переключается на воспроизведение из сегмента 9 (принимаемого через широковещательную передачу).

Этот вариант использования демонстрирует "быструю настройку" с помощью одноадресной передачи. В этом случае, поставщик услуг хочет распространять одно представление, которое имеет высокую SAR-частоту (типично, тип 3 является возможным) для быстрого доступа. Тем не менее, после настройки, клиент хочет переключаться на представление, которое является более эффективным и которое имеет меньше IDR-кадров. Представление, на которое выполняется переключение, может даже иметь другой размер сегмента. Этот сценарий может иметь место в одноадресной передаче, но также и в гибридном случае. Сценарий показан на фиг. 1. На этой схеме, более короткие сегменты заданы доступными через одноадресную передачу, причем каждый сегмент включает в себя IDR-кадр. Если клиент присоединяется к программе в определенное время и без поддержки одноадресной передачи, требуется некоторое время до тех пор, пока сегмент не примется и сможет начинать воспроизводиться (сегмент 9 на фиг. 1). Это обусловлено тем фактом, что должен приниматься полный сегмент (чтобы надлежащим образом инициализировать, например, декодер мультимедиа с возможностью декодировать мультимедийные данные сегмента).

В этом случае, представление одноадресной передачи предлагается с четвертью длительности сегмента. Клиент может сразу выбирать воспроизводить одноадресные короткие сегменты до тех пор, пока эффективное (длинный сегмент, больше расстояние между IDR-кадрами) представление широковещательной передачи не поступит через широковещательную передачу. Передача в сигналах этих характеристик (позиции точек произвольного доступа и точек переключения) в MPD является релевантной, но невозможной на сегодня.

Другой аналогичный вариант использования включает в себе быструю настройку с помощью SHVC. Может быть предусмотрено предложение базового слоя с низкой RAP-частотой и даже низким размером сегмента и улучшающий слой, который имеет большую GOP-частоту. В таком случае означенное должно достигаться, как пояснено относительно фиг. 1. На сегодня передача в сигналах этих признаков является невозможной.

Другой желательный вариант использования заключается в использовании эффективного буфера сдвига во времени. В определенных случаях, представление может предлагаться на границе передач вживую с небольшими сегментами, но как только клиент перемещается в буфер сдвига во времени, размер сегмента увеличивается. Представления по-прежнему должны находиться в одном адаптивном наборе, чтобы выражать характеристики прозрачного переключения, но они не должны принудительно иметь идентичные размеры сегментов и/или идентичную частоту точек переключения/точек произвольного доступа. То же применимо для записи передаваемого вживую события для будущего использования по запросу.

Другой вариант использования включает в себе быструю настройку с помощью открытых GOP. Открытая GOP, в общем, может соответствовать GOP, включающей в себя изображения, которые могут прогнозироваться относительно изображений за пределами GOP. В этом состоит отличие от закрытой

GOP, которая является автономной в том, что все изображения GOP прогнозируются из других изображений в пределах GOP. Например, открытая GOP может начинаться со взаимно прогнозированного изображения (или взаимно прогнозированного ключевого кадра), тогда как закрытая GOP может начинаться с внутренне прогнозированного изображения.

Случай быстрой настройки с помощью открытых GOP может быть типичным случаем для широко-вещательной быстрой настройки. Проблема заключается в том, что возникают случаи, в которых требуется быстро настраиваться, переключаться между представлениями и возможно обеспечивать низкую задержку. Это может приводить к сложным вариантам использования для передачи сигналов, а именно, передачи в сигналах сегментов, открытых GOP, закрытых GOP, совмещений сегментов и т.д.

Другой вариант использования включает в себе быстрое понижающее переключение для непрерывности. Этот случай также может быть типичным для сценария широковещательной быстрой настройки. Проблема заключается в том, что возникают случаи, для которых требуется быстро настраиваться, переключаться между представлениями и возможно обеспечивать низкую задержку. Это может приводить к сложным вариантам использования для передачи сигналов, а именно, передачи в сигналах сегментов, открытых GOP, закрытых GOP, совмещений сегментов и т.д.

Другой вариант использования включает в себе доступности сегментов. Чтобы уменьшать времена задержки, мало того, что сегменты должны быть короткими, но также и время между формированием сегментов и публикацией должно быть небольшим. Чтобы не допускать ошибок HTTP 404, времена доступности сегментов должны быть доступными (например, передаваться в сигналах) в приемное устройство. Шаблоны сегментов обеспечивают шаблон, чтобы уведомлять в отношении времен доступности, но это требует того, что сегменты должны быть доступны в то точное время, и в силу этого варьирования длительностей сегментов должны учитываться при уведомлении в отношении начальных времен доступности сегментов, и кодер должен придерживаться этого шаблона. Если поставщик контента не вынуждается формировать IDR-кадр с временами доступности сегментов, он может проще варьировать размещения IDR-кадра, и времена доступности сегментов могут уведомляться точнее. Этот аспект должен рассматриваться при передаче в сигналах длительностей сегментов.

В различных вариантах использования, различные признаки переключения, доставки и произвольного доступа являются более или менее релевантными, но они, возможно, должны обеспечиваться в одном предложении контента. Существуют несколько сценариев, которые должны рассматриваться:

Развертывание широковещательного распространения с малым временем получения каналов наряду со способностью переключаться на представление одноадресной передачи на более низкой частоте.

Доставка версии с низкой задержкой на границе передач вживую по одноадресной передаче, которая синхронизируется с широковещательной передачей.

Доставка версии с низкой задержкой по широковещательной передаче только с большей частотой произвольного доступа, чем единицы доставки.

Переменные длительности сегментов, которые должны учитываться.

Технологии этого раскрытия сущности могут обеспечивать возможность этих различных вариантов использования, отдельно или в любой комбинации, и могут преодолевать любые из проблем, поясненных выше.

Фиг. 2 является диаграммой 200 ВЕННА, иллюстрирующей взаимосвязи между различными типами мультимедийных сегментов. Мультимедийные сегменты могут использоваться для любых различных целей в DASH, к примеру, для следующего:

Переключение представлений

Закрытые GOP, в общем, требуются;

Сегменты не должны перекрываться во времени в пределах одного представления;

Сегменты должны совмещаться между различными представлениями в одном адаптивном наборе.

Произвольный доступ

Открытая GOP, в общем, требуется;

Сегменты могут перекрываться во времени в пределах одного представления, если разрешается произвольный доступ открытой GOP.

Единица доставки

Отсутствуют требования по произвольному доступу или переключению;

Сегмент должен включать в себя целое число кинофрагментов.

Чтобы разрешать различные аспекты, четыре различных типа (или формата) сегментов могут рассматриваться согласно фиг. 2:

Формат 202 сегментов единиц доставки: Только фрагмент без ограничений. (Представлен посредством эллипса со сплошным контуром на фиг. 2);

Формат 204 сегментов с произвольным доступом: Открытая GOP для настройки. (Представлен посредством эллипса с пунктирным контуром на фиг. 2);

Формат 206 сегментов без перекрытия: Клиентское устройство может переключаться на сегмент этого формата без каких-либо проблем. (Представлен посредством эллипса с пунктирным контуром на фиг. 2.);

Формат 208 сегментов с переключением: Клиентское устройство может переключаться на сегмент этого формата. (Представлен посредством эллипса с пунктирным контуром с двумя точками на фиг. 2.).

Фиг. 3 является концептуальной схемой, иллюстрирующей примерную структуру представления 210 и ISO BMFF-файлов 212A-212C. фиг. 3 также показывает покомпонентный вид ISO BMFF-файла 212A, который включает в себя поле moof (кинофрагментов) и поле киноданных (mdat). Примерный ISO BMFF-файл 212A по фиг. 3 концептуально является аналогичным кинофрагментам 164 по фиг. 6, подробнее описанным ниже. Релевантно считать, что кинофрагменты представляют собой единицы доставки для мультимедийных данных. Кинофрагменты формируются таким образом, что они содержат последовательность поля moof и поля mdat, например, как показано на фиг. 3.

Фиг. 4 является блок-схемой, иллюстрирующей примерную систему 10, которая реализует технологии для потоковой передачи мультимедийных данных по сети. В этом примере, система 10 включает в себя устройство 20 подготовки контента, серверное устройство 60 и клиентское устройство 40. Клиентское устройство 40 и серверное устройство 60 функционально соединяются посредством сети 74, которая может содержать Интернет. В некоторых примерах, устройство 20 подготовки контента и серверное устройство 60 также могут соединиться посредством сети 74 или другой сети либо могут непосредственно функционально соединиться. В некоторых примерах, устройство 20 подготовки контента и серверное устройство 60 могут содержать идентичное устройство.

Устройство 20 подготовки контента, в примере по фиг. 4, содержит аудиоисточник 22 и видеоисточник 24. Аудиоисточник 22 может содержать, например, микрофон, который формирует электрические сигналы, представляющие захваченные аудиоданные, которые должны кодироваться посредством аудиокодера 26. Альтернативно, аудиоисточник 22 может содержать носитель хранения данных, сохраняющий ранее записанные аудиоданные, формирователь аудиоданных, такой как компьютеризированный синтезатор, или любой другой источник аудиоданных. Видеоисточник 24 может содержать видеокамеру, которая формирует видеоданные, которые должны кодироваться посредством видеокодера 28, носитель хранения данных, кодированный с помощью ранее записанных видеоданных, модуль формирования видеоданных, такой как источник компьютерной графики, или любой другой источник видеоданных. Устройство 20 подготовки контента не обязательно функционально соединяется с серверным устройством 60 во всех примерах, но может сохранять мультимедийный контент на отдельный носитель, который считывается посредством серверного устройства 60.

Необработанные аудио- и видеоданные могут содержать аналоговые или цифровые данные. Аналоговые данные могут быть оцифрованы до кодирования посредством аудиокодера 26 и/или видеокодера 28. Аудиоисточник 22 может получать аудиоданные от говорящего участника в то время, когда говорящий участник говорит, и видеоисточник 24 может одновременно получать видеоданные говорящего участника. В других примерах, аудиоисточник 22 может содержать машиночитаемый носитель хранения данных, содержащий сохраненные аудиоданные, и видеоисточник 24 может содержать машиночитаемый носитель хранения данных, содержащий сохраненные видеоданные. Таким образом, технологии, описанные в этом раскрытии сущности, могут применяться потоковым аудио- и видеоданным в реальном времени передач вживую либо к заархивированным, предварительно записанным аудио- и видеоданным.

Аудиокадры, которые соответствуют видеокдрам, в общем, представляют собой аудиокадры, содержащие аудиоданные, который захвачены (или сформированы) посредством аудиоисточника 22 одновременно с видеоданными, захваченными (или сформированными) посредством видеоисточника 24, который содержится в видеокдрах. Например, в то время когда говорящий участник, в общем, формирует аудиоданные посредством разговора, аудиоисточник 22 захватывает аудиоданные, а видеоисточник 24 захватывает видеоданные говорящего участника одновременно, т.е. в то время когда аудиоисточник 22 захватывает аудиоданные. Следовательно, аудиокадр может временно соответствовать одному или более конкретным видеокдрам. Соответственно, аудиокадр, соответствующий видеокдру, в общем, соответствует ситуации, в которой аудиоданные и видеоданные захвачены одновременно, и для которой аудиокадр и видеокдр содержат, соответственно, аудиоданные и видеоданные, которые захвачены одновременно.

В некоторых примерах, аудиокодер 26 может кодировать временную метку в каждом кодированном аудиокадре, который представляет время, в которое записаны аудиоданные для кодированного аудиокадра, и аналогично, видеокодер 28 может кодировать временную метку в каждом кодированном видеокдре, который представляет время, в которое записаны видеоданные для кодированного видеокдра. В таких примерах, аудиокадр, соответствующий видеокдру, может содержать аудиокадр, содержащий временную метку, и видеокдр, содержащий идентичную временную метку. Устройство 20 подготовки контента может включать в себя внутренний тактовый генератор, из которого аудиокодер 26 и/или видеокодер 28 могут формировать временные метки, или который аудиоисточник 22, и видеоисточник 24 могут использовать для того, чтобы ассоциировать аудио- и видеоданные, соответственно, с временной меткой.

В некоторых примерах, аудиоисточник 22 может отправлять данные в аудиокодер 26 согласно времени, в которое записаны видеоданные, и видеоисточник 24 может отправлять данные в видеокодер 28 согласно времени, в которое записаны видеоданные. В некоторых примерах, аудиокодер 26 может коди-

ровать идентификатор последовательности в кодированных аудиоданных, чтобы указывать относительное временное упорядочение кодированных аудиоданных, но без обязательного указания абсолютного времени, в которое записаны видеоданные, и аналогично, видеокодер 28 также может использовать идентификаторы последовательностей, чтобы указывать относительное временное упорядочение кодированных видеоданных. Аналогично, в некоторых примерах, идентификатор последовательности может преобразовываться или иным образом коррелироваться с временной меткой.

Аудиокодер 26, в общем, формирует поток кодированных аудиоданных, тогда как видеокодер 28 формирует поток кодированных видеоданных. Каждый отдельный поток данных (аудио- или видео-) может упоминаться как элементарный поток. Элементарный поток представляет собой один кодированный в цифровой форме (возможно сжатый) компонент представления. Например, кодированная видео- или аудиочасть представления может представлять собой элементарный поток. Элементарный поток может преобразовываться в пакетизированный элементарный поток (PES) перед инкапсуляцией в видео-файле. В идентичном представлении, идентификатор потока может использоваться для того, чтобы отличать PES-пакеты, принадлежащие одному элементарному потоку, от других. Базовая единица данных элементарного потока представляет собой пакет пакетизированных элементарных потоков (PES). Таким образом, кодированные видеоданные, в общем, соответствуют элементарным видеопотокам. Аналогично, аудиоданные соответствуют одному или более соответствующих элементарных потоков.

Множество стандартов кодирования видео, таких как ITU-T H.264/AVC и стандарт высокоэффективного кодирования видео (HEVC), задают синтаксис, семантику и процессы декодирования для безошибочных потоков битов, любой из которых соответствует определенному профилю или уровню. Стандарты кодирования видео типично не указывают кодер, но кодер имеет задачу гарантирования того, что сформированные потоки битов являются совместимыми со стандартом для декодера. В контексте стандартов кодирования видео, "профиль" соответствует поднабору алгоритмов, признаков или инструментальных средств и ограничений, которые применяются к ним. Как задано посредством стандарта H.264, например, "профиль" представляет собой поднабор полного синтаксиса потока битов, который указывается посредством стандарта H.264. "Уровень" соответствует ограничениям по потреблению ресурсов декодера, таких как запоминающее устройство декодера и вычисление, которые связаны с разрешением изображений, скорости передачи битов и скорость блочной обработки. Профиль может передаваться в сигналах со значением `profile_idc` (индикатора профиля), тогда как уровень может передаваться в сигналах со значением `level_idc` (индикатора уровня).

Стандарт H.264, например, распознает то, что в пределах, налагаемых посредством синтаксиса данного профиля, по-прежнему можно требовать большого варьирования производительности кодеров и декодеров в зависимости от значений, принимаемых посредством синтаксических элементов в потоке битов, таких как указанный размер декодированных изображений. Стандарт H.264 дополнительно распознает, что во многих вариантах применения, непрактично и неэкономично реализовывать декодер, допускающий затрагивание всех гипотетических вариантов использования синтаксиса в конкретном профиле. Соответственно, стандарт H.264 задает "уровень" в качестве указанного набора ограничений, налагаемых на значения синтаксических элементов в потоке битов. Эти ограничения могут представлять собой простые пределы для значений. Альтернативно, эти ограничения могут принимать форму ограничений на арифметические комбинации значений (например, ширина изображения, умноженная на высоту изображения, умноженную на число изображений, декодированных в секунду). Стандарт H.264 дополнительно обеспечивает то, что отдельные реализации могут поддерживать различный уровень для каждого поддерживаемого профиля.

Декодер, соответствующий профилю, обычно поддерживает все функции, заданные в профиле. Например, в качестве признака кодирования, кодирование В-изображений не поддерживается в базовом профиле H.264/AVC, но поддерживается в других профилях H.264/AVC. Декодер, соответствующий уровню, должен допускать декодирование любого потока битов, который не требует ресурсов за рамками ограничений, заданных в уровне. Определения профилей и уровней могут быть полезными для интерпретируемости. Например, в ходе передачи видео, пара определений профиля и уровня может быть согласована и установлена для всего сеанса передачи. Более конкретно, в H.264/AVC, уровень может задавать ограничения на число макроблоков, которые должны обрабатываться, размер буфера декодированных изображений (DPB), размер буфера кодированных изображений (CPB), вертикальный диапазон векторов движения, максимальное число векторов движения в расчете на два последовательных MB, а также то, может или нет блок В иметь меньшее число сегментов субмакроблока, чем 8×8 пикселей. Таким образом, декодер может определять то, допускает или нет декодер надлежащее декодирование потока битов.

В примере по фиг. 4, модуль 30 инкапсуляции устройства 20 подготовки контента принимает элементарные потоки, содержащие кодированные видеоданные, из видеокодера 28 и элементарные потоки, содержащие кодированные аудиоданные, из аудиокодера 26. В некоторых примерах, видеокодер 28 и аудиокодер 26 могут включать в себя модули пакетизации для формирования PES-пакетов из кодированных данных. В других примерах, видеокодер 28 и аудиокодер 26 могут взаимодействовать с соответствующими модулями пакетизации для формирования PES-пакетов из кодированных данных. В еще одних

других примерах, модуль 30 инкапсуляции может включать в себя модули пакетизации для формирования PES-пакетов из кодированных аудио- и видеоданных.

Видеокодер 28 может кодировать видеоданные мультимедийного контента множеством способов, чтобы формировать различные представления мультимедийного контента на различных скоростях передачи битов и с различными характеристиками, такими как разрешения в пикселях, частоты кадров, соответствие различным стандартам кодирования, соответствие различным профилям и/или уровням профилей для различных стандартов кодирования, представления, имеющие один или более видов (например, для двумерного или трехмерного воспроизведения), или с другими такими характеристиками. Представление, при использовании в этом раскрытии сущности, может содержать одно из аудиоданных, видеоданных, текстовых данных (например, для скрытых титров) или других таких данных. Представление может включать в себя элементарный поток, к примеру, элементарный аудиопоток или элементарный видеопоток. Каждый PES-пакет может включать в себя stream_id, который идентифицирует элементарный поток, которому принадлежит PES-пакет. Модуль 30 инкапсуляции отвечает за ассемблирование элементарных потоков в видеофайлы (например, сегменты) различных представлений.

Модуль 30 инкапсуляции принимает PES-пакеты для элементарных потоков представления из аудиокодера 26 и видеокодера 28 и формирует соответствующие единицы уровня абстрагирования от сети (NAL) из PES-пакетов. В примере H.264/AVC (усовершенствованное кодирование видео), кодированные видеосегменты организуются в NAL-единицы, которые обеспечивают "оптимизированное для использования в сети" видеопредставление, нацеленное на такие варианты применения, как видеотелефония, хранение данных, широковещательная передача или потоковая передача. NAL-единицы могут классифицироваться на NAL-единицы уровня кодирования видео (VCL) и не-VCL NAL-единицы. VCL-единицы могут содержать базовый механизм сжатия и могут включать в себя данные уровня блока, макроблока и/или серии последовательных макроблоков. Другие NAL-единицы могут представлять собой не-VCL NAL-единицы. В некоторых примерах, кодированное изображение в один момент времени, нормально представленное в качестве первичного кодированного изображения, может содержаться в единице доступа, которая может включать в себя одну или более NAL-единиц.

Не-VCL NAL-единицы могут включать в себя NAL-единицы наборов параметров и SEI NAL-единицы, в числе прочего. Наборы параметров могут содержать информацию заголовка уровня последовательности (в наборах параметров последовательности (SPS)) и нечасто изменяющуюся информацию заголовка уровня изображения (в наборах параметров изображения (PPS)). Для наборов параметров (например, PPS и SPS), нечасто изменяющаяся информация не должна повторяться для каждой последовательности или изображения, и как следствие, может повышаться эффективность кодирования. Кроме того, использование наборов параметров может обеспечивать внеполосную передачу важной информации заголовка, исключая необходимость избыточных передач для обеспечения устойчивости к ошибкам. В примерах внеполосной передачи, NAL-единицы наборов параметров могут передаваться по каналу, отличному от канала для других NAL-единиц, к примеру, SEI NAL-единиц.

Дополнительная улучшающая информация (SEI) может содержать информацию, которая не требуется для декодирования выборок кодированных изображений из VCL NAL-единиц, но может помогать в процессах, связанных с декодированием, отображением, устойчивостью к ошибкам и другими целями. SEI-сообщения могут содержаться в не-VCL NAL-единицах. SEI-сообщения представляют собой нормативную часть некоторых стандартных технических требований и в силу этого не всегда являются обязательными для совместимой со стандартом реализации декодера. SEI-сообщения могут представлять собой SEI-сообщения уровня последовательности или SEI-сообщения уровня изображения. Некоторая информация уровня последовательности может содержаться в SEI-сообщениях, к примеру, в SEI-сообщениях с информацией масштабируемости в примере SVC и SEI-сообщениях с информацией масштабируемости видов в MVC. Эти примерные SEI-сообщения могут передавать информацию относительно, например, извлечения рабочих точек и характеристик рабочих точек. Помимо этого, модуль 30 инкапсуляции может формировать файл манифеста, такой как дескриптор мультимедийного представления (MPD), который описывает характеристики представлений. Модуль 30 инкапсуляции может форматировать MPD согласно расширяемому языку разметки (XML).

Модуль 30 инкапсуляции может обеспечивать данные для одного или более представлений мультимедийного контента, вместе с файлом манифеста (например, MPD), в интерфейс 32 вывода. Интерфейс 32 вывода может содержать сетевой интерфейс или интерфейс для записи на носитель хранения данных, к примеру, интерфейс универсальной последовательной шины (USB), устройство или средство записи CD или DVD, интерфейс с магнитными носителями хранения данных или носителями хранения данных на основе флэш-либо либо другие интерфейсы для хранения или передачи мультимедийных данных. Модуль 30 инкапсуляции может обеспечивать данные каждого из представлений мультимедийного контента в интерфейс 32 вывода, который может отправлять данные в серверное устройство 60 через сетевую передачу или носители хранения данных. В примере по фиг. 4, серверное устройство 60 включает в себя носитель 62 хранения данных, который сохраняет различный мультимедийный контент 64, включающий в себя соответствующий файл 66 манифеста и одно или более представлений 68A-68N (представлений 68). В некоторых примерах, интерфейс 32 вывода также может отправлять данные непосред-

ственно в сеть 74.

В некоторых примерах, представления 68 могут разделяться на адаптивные наборы. Иными словами, различные поднаборы представлений 68 могут включать в себя соответствующие общие наборы характеристик, такие как кодек, профиль и уровень, разрешение, число видов, формат файлов для сегментов, информация типов текста, которая может идентифицировать язык или другие характеристики текста, которые должны отображаться с представлением, и/или аудиоданные, которые должны декодироваться и представляться, например, посредством динамиков, информация ракурсов камеры, которая может описывать ракурс камеры или перспективу сцены в камере реального мира для представлений в адаптивном наборе, рейтинговая информация, которая описывает пригодность контента для конкретных зрителей и т.п.

Файл 66 манифеста может включать в себя данные, указывающие поднаборы представлений 68, соответствующих конкретным адаптивным наборам, а также общие характеристики для адаптивных наборов. Файл 66 манифеста также может включать в себя данные, представляющие отдельные характеристики, такие как скорости передачи битов, для отдельных представлений адаптивных наборов. Таким образом, адаптивный набор может обеспечивать упрощенную адаптацию полосы пропускания сети. Представления в адаптивном наборе могут указываться с использованием дочерних элементов для элемента адаптивного набора файла 66 манифеста.

Серверное устройство 60 включает в себя модуль 70 обработки запросов и сетевой интерфейс 72. В некоторых примерах, серверное устройство 60 может включать в себя множество сетевых интерфейсов. Кроме того, любые из признаков серверного устройства 60 могут реализовываться на других устройствах сети доставки контента, таких как маршрутизаторы, мосты, прокси-устройства, коммутаторы или другие устройства. В некоторых примерах, промежуточные устройства сети доставки контента могут кэшировать данные мультимедийного контента 64 и включать в себя компоненты, которые фактически соответствуют компонентам серверного устройства 60. В общем, сетевой интерфейс 72 выполнен с возможностью отправлять и принимать данные через сеть 74.

Модуль 70 обработки запросов выполнен с возможностью принимать сетевые запросы из клиентских устройств, таких как клиентское устройство 40, на предмет данных носителя 62 хранения данных. Например, модуль 70 обработки запросов может реализовывать версию 1.1 протокола передачи гипертекста (HTTP), как описано в документе RFC 2616 "Hypertext Transfer Protocol -HTTP/1.1" авторов R. Fielding и др., Network Working Group, IETF, июнь 1999 года. Иными словами, модуль 70 обработки запросов может быть выполнен с возможностью принимать HTTP GET-или частичные GET-запросы и обеспечивать данные мультимедийного контента 64 в ответ на запросы. Запросы могут указывать сегмент одного из представлений 68, например, с использованием URL-адреса сегмента. В некоторых примерах, запросы также могут указывать один или более байтовых диапазонов сегмента, за счет этого включая в себя частичные GET-запросы. Модуль 70 обработки запросов дополнительно может быть выполнен с возможностью обслуживать HTTP HEAD-запросы, чтобы обеспечивать данные заголовка сегмента одного из представлений 68. В любом случае, модуль 70 обработки запросов может быть выполнен с возможностью обрабатывать запросы, чтобы обеспечивать запрашиваемые данные в запрашивающее устройство, к примеру, в клиентское устройство 40.

Дополнительно или альтернативно, модуль 70 обработки запросов может быть выполнен с возможностью доставлять мультимедийные данные через протокол ширококвещательной или многоадресной передачи, такой как eMBMS. Устройство 20 подготовки контента может создавать DASH-сегменты и/или субсегменты практически идентично тому, что описано, но серверное устройство 60 может доставлять эти сегменты или субсегменты с использованием eMBMS либо другого сетевого транспортного протокола ширококвещательной или многоадресной передачи. Например, модуль 70 обработки запросов может быть выполнен с возможностью принимать запрос на присоединение к группе многоадресной передачи из клиентского устройства 40. Иными словами, серверное устройство 60 может оповещать адрес по Интернет-протоколу (IP), ассоциированный с группой многоадресной передачи, в клиентские устройства, включающие в себя клиентское устройство 40, ассоциированное с конкретным мультимедийным контентом (например, ширококвещательной передачей передаваемого вживую события). Клиентское устройство 40, в свою очередь, может отправлять запрос на то, чтобы присоединиться к группе многоадресной передачи. Этот запрос может распространяться по сети 74, например, через маршрутизаторы, составляющие сеть 74, так что маршрутизаторам инструктируется направлять трафик, предназначенный для IP-адреса, ассоциированного с группой многоадресной передачи, в подписанные клиентские устройства, такие как клиентское устройство 40.

Как проиллюстрировано в примере по фиг. 4, мультимедийный контент 64 включает в себя файл 66 манифеста, который может соответствовать описанию мультимедийного представления (MPD). Файл 66 манифеста может содержать описания различных альтернативных представлений 68 (например, услуг передачи видео с различным качеством), и описание может включать в себя, например, информацию кодека, значение профиля, значение уровня, скорость передачи битов и другие описательные характеристики представлений 68. Клиентское устройство 40 может извлекать MPD мультимедийного представления, чтобы определять то, как осуществлять доступ к сегментам представлений 68.

В частности, модуль 52 извлечения может извлекать конфигурационные данные (не показаны) клиентского устройства 40, чтобы определять характеристики декодирования видеodeкодера 48 и характеристики рендеринга видеовывода 44. Конфигурационные данные также могут включать в себя любое из настройки языка, выбранной пользователем клиентского устройства 40, одной или более перспектив камеры, соответствующих настройкам глубины, заданным пользователем клиентского устройства 40, и/или настройки рейтинга, выбранной пользователем клиентского устройства 40. Модуль 52 извлечения может содержать, например, веб-обозреватель или мультимедийный клиент, выполненный с возможностью отправлять HTTP GET и частичные GET-запросы. Модуль 52 извлечения может соответствовать программным инструкциям, выполняемым посредством одних или более процессоров или модулей обработки (не показаны) клиентского устройства 40. В некоторых примерах, вся или части функциональности, описанной относительно модуля 52 извлечения, может реализовываться в аппаратных средствах или комбинации аппаратных средств, программного обеспечения и/или микропрограммного обеспечения, при этом требуемые аппаратные средства могут обеспечиваться, чтобы выполнять инструкции для программного обеспечения или микропрограммного обеспечения.

Модуль 52 извлечения может сравнивать характеристики декодирования и рендеринга клиентского устройства 40 с характеристиками представлений 68, указываемыми посредством информации файла 66 манифеста. Модуль 52 извлечения может первоначально извлекать, по меньшей мере, часть файла 66 манифеста, чтобы определять характеристики представлений 68. Например, модуль 52 извлечения может запрашивать часть файла 66 манифеста, которая описывает характеристики одного или более адаптивных наборов. Модуль 52 извлечения может выбирать поднабор представлений 68 (например, адаптивный набор), имеющий характеристики, которые могут удовлетворяться посредством характеристик кодирования и рендеринга клиентского устройства 40. Модуль 52 извлечения затем может определять скорости передачи битов для представлений в адаптивном наборе, определять доступную на сегодня величину полосы пропускания сети и извлекать сегменты из одного из представлений, имеющих скорость передачи битов, которая может удовлетворяться посредством полосы пропускания сети.

В общем, представления с более высокой скоростью передачи битов могут давать в результате воспроизведение видео более высокого качества, тогда как представления с более низкой скоростью передачи битов могут обеспечивать воспроизведение видео достаточного качества, когда доступная полоса пропускания сети снижается. Соответственно, когда доступная полоса пропускания сети является относительно высокой, модуль 52 извлечения может извлекать данные из представлений с относительно высокой скоростью передачи битов, тогда как когда доступная полоса пропускания сети является низкой, модуль 52 извлечения может извлекать данные из представлений с относительно низкой скоростью передачи битов. Таким образом, клиентское устройство 40 может передавать в потоковом режиме мультимедийные данные по сети 74 при одновременной адаптации к изменяющейся доступности полосы пропускания сети для сети 74.

Дополнительно или альтернативно, модуль 52 извлечения может быть выполнен с возможностью принимать данные в соответствии с сетевым протоколом широкоэмитальной или многоадресной передачи, таким как eMBMS или многоадресная IP-передача. В таких примерах, модуль 52 извлечения может отправлять запрос, чтобы присоединиться к сетевой группе многоадресной передачи, ассоциированной с конкретным мультимедийным контентом. После присоединения к группе многоадресной передачи, модуль 52 извлечения может принимать данные группы многоадресной передачи без дополнительных запросов, выданных в серверное устройство 60 или устройство 20 подготовки контента. Модуль 52 извлечения может отправлять запрос, чтобы выходить из группы многоадресной передачи, когда данные группы многоадресной передачи более не требуются, например, чтобы прекращать воспроизведение или переключать каналы на другую группу многоадресной передачи.

Сетевой интерфейс 54 может принимать и обеспечивать данные сегментов выбранного представления в модуль 52 извлечения, который, в свою очередь, может обеспечивать сегменты в модуль 50 декапсуляции. Модуль 50 декапсуляции может декапсулировать элементы видеофайла на составляющие PES-потоки, депакетизировать PES-потоки, чтобы извлекать кодированные данные, и отправлять кодированные данные в аудиodeкодер 46 или в видеodeкодер 48, в зависимости от того, являются кодированные данные частью аудио- или видеопотока, например, как указано посредством заголовков PES-пакетов потока. Аудиodeкодер 46 декодирует кодированные аудиоданные и отправляет декодированные аудиоданные в аудиовывод 42, тогда как видеodeкодер 48 декодирует кодированные видеоданные и отправляет декодированные видеоданные, которые могут включать в себя множество видов потока, в видеовывод 44.

В соответствии с технологиями этого раскрытия сущности, файл 66 манифеста может передавать в сигналах различные форматы сегментов, которым могут соответствовать сегменты (также упоминаются в данном документе в качестве типов сегментов). Файл 66 манифеста также может передавать в сигналах местоположения сегментов, которые соответствуют каждому формату (т.е. местоположения каждого из различных типов сегментов). Например, файл 66 манифеста может передавать в сигналах частоты, на которых каждый из различных типов сегментов возникает в каждом из представлений 68.

С использованием файла 66 манифеста, клиентское устройство 40 может достигать воспроизведе-

ния с низкой задержкой мультимедийных данных. Например, одно из представлений 68 (например, представление 68A) может включать в себя SAP на относительно высокой частоте, как указано посредством файла 66 манифеста, тогда как другое из представлений 68 (например, представление 68N) может включать в себя SAP на относительно низкой частоте. В частности, SAP могут составлять часть сегментов, соответствующих конкретным форматам, например, формату мультимедийных сегментов с произвольным доступом и/или формату мультимедийных сегментов с переключением. Кроме того, представления 68 могут быть доступными для извлечения через различные услуги передачи. Например, представление 68A может быть доступным через одноадресную передачу, тогда как представление 68N может быть доступным через широковещательную передачу.

В соответствии с некоторыми примерами технологий этого раскрытия сущности, клиентское устройство 40 может определять, согласно вышеприведенному примеру, что представление 68A включает в себя относительно высокую частоту SAP (например, очень частые мультимедийные сегменты с произвольным доступом и/или очень частые мультимедийные сегменты с переключением), как указано посредством файла 66 манифеста. Кроме того, клиентское устройство 40 может определять то, что представление 68N включает в себя относительно низкую частоту SAP, но также и имеет относительно более высокое качество. Таким образом, чтобы инициировать извлечение мультимедийных данных, клиентское устройство 40 может начинать посредством извлечения мультимедийных сегментов из представления 68A до тех пор, пока клиентское устройство 40 не сможет переключиться на представление 68N, например, в мультимедийном сегменте с произвольным доступом или мультимедийном сегменте с переключением из 68N, как указано посредством файла 66 манифеста. Различные подробные варианты использования, описывающие примеры этих технологий, описываются ниже относительно, например, фиг. 7-14.

Видеокодер 28, декодер 48, аудиокодер 26, декодер 46, модуль 30 инкапсуляции, модуль 52 извлечения и модуль 50 декапсуляции могут реализовываться как любая из множества подходящих схем обработки, при соответствующих условиях, к примеру, как один или более микропроцессоров, процессоров цифровых сигналов (DSP), специализированных интегральных схем (ASIC), программируемых пользователем вентильных матриц (FPGA), дискретная логическая схема, программное обеспечение, аппаратные средства, микропрограммное обеспечение либо любые комбинации вышеозначенного. Каждый из видеокодера 28 и декодера 48 может быть включен в один или более кодировщиков или декодеров, любой из которых может быть интегрирован как часть комбинированного видеокодера/декодера (кодека). Аналогично, каждый из аудиокодера 26 и декодера 46 может быть включен в один или более кодировщиков или декодеров, любой из которых может быть интегрирован как часть комбинированного кодека. Устройство, включающее в себя видеокодер 28, декодер 48, аудиокодер 26, декодер 46, модуль 30 инкапсуляции, модуль 52 извлечения и/или модуль 50 декапсуляции, может содержать интегральную схему, микропроцессор и/или устройство беспроводной связи, такое как сотовый телефон.

Клиентское устройство 40, серверное устройство 60 и/или устройство 20 подготовки контента могут быть выполнены с возможностью работать в соответствии с технологиями этого раскрытия сущности. В целях примера, это раскрытие сущности описывает эти технологии относительно клиентского устройства 40 и серверного устройства 60. Тем не менее, следует понимать, что устройство 20 подготовки контента может быть выполнено с возможностью осуществлять эти технологии, вместо (или в дополнение к) серверного устройства 60.

Модуль 30 инкапсуляции может формировать NAL-единицы, содержащие заголовок, который идентифицирует программу, которой принадлежит NAL-единица, а также рабочие данные, например, аудиоданные, видеоданные или данные, которые описывают транспортный или программный поток, которому соответствует NAL-единица. Например, в H.264/AVC, NAL-единица включает в себя 1-байтовый заголовок и рабочие данные варьирующегося размера. NAL-единица, включающая в себя видеоданные в своих рабочих данных, может содержать различные уровни детализации видеоданных. Например, NAL-единица может содержать блок видеоданных, множество блоков, серию последовательных макроблоков видеоданных или полное изображение видеоданных. Модуль 30 инкапсуляции может принимать кодированные видеоданные из видеокодера 28 в форме PES-пакетов элементарных потоков. Модуль 30 инкапсуляции может ассоциировать каждый элементарный поток с соответствующей программой.

Модуль 30 инкапсуляции также может ассемблировать единицы доступа из множества NAL-единиц. В общем, единица доступа может содержать одну или более NAL-единиц для представления кадра видеоданных, также аудиоданных, соответствующих кадру, когда такие аудиоданные доступны. Единица доступа, в общем, включает в себя все NAL-единицы для одного момента времени вывода, например, все аудио- и видеоданные для одного момента времени. Например, если каждый вид имеет частоту кадров в 20 кадров в секунду (кадр/с), то каждый момент времени может соответствовать временному интервалу в 0,05 секунд. В течение этого временного интервала, конкретные кадры для всех видов идентичной единицы доступа (идентичного момента времени) могут подготавливаться посредством рендеринга одновременно. В одном примере, единица доступа может содержать кодированное изображение в один момент времени, которое может представляться в качестве первичного кодированного изображения.

Соответственно, единица доступа может содержать все аудио- и видеокдры общего временного

момента, например, все виды, соответствующие времени X. Это раскрытие сущности также упоминает кодированное изображение конкретного вида в качестве "компонента вида". Иными словами, компонент вида может содержать кодированное изображение (или кадр) для конкретного вида в конкретное время. Соответственно, единица доступа может задаваться как включающая в себя все компоненты видов общего временного момента. Порядок декодирования единиц доступа необязательно должен быть идентичным порядку вывода или отображения.

Мультимедийное представление может включать в себя описание мультимедийного представления (MPD), которое может содержать описания различных альтернативных представлений (например, услуг передачи видео с различным качеством), и описание может включать в себя, например, информацию кодека, значение профиля и значение уровня. MPD является одним примером файла манифеста, такого как файл 66 манифеста. Клиентское устройство 40 может извлекать MPD мультимедийного представления, чтобы определять то, как осуществлять доступ к кинофрагментам различных представлений. Кинофрагменты могут быть расположены в полях кинофрагментов (полях *moof*) видеофайлов.

Файл 66 манифеста (который может содержать, например, MPD) может оповещать доступность сегментов представлений 68. Иными словами, MPD может включать в себя информацию, указывающую физическое время, в которое первый сегмент одного из представлений 68 становится доступным, а также информацию, указывающую длительность сегментов в представлениях 68. Таким образом, модуль 52 извлечения клиентского устройства 40 может определять то, когда каждый сегмент доступен, на основе начального времени, а также длительности сегментов, предшествующих конкретному сегменту.

После того, как модуль 30 инкапсуляции ассемблирует NAL-единицы и/или единицы доступа в видеофайл на основе принимаемых данных, модуль 30 инкапсуляции передает видеофайл в интерфейс 32 вывода для вывода. В некоторых примерах, модуль 30 инкапсуляции может сохранять видеофайл локально или отправлять видеофайл на удаленный сервер через интерфейс 32 вывода, вместо отправки видеофайла непосредственно в клиентское устройство 40. Интерфейс 32 вывода может содержать, например, передающее устройство, приемо-передающее устройство, устройство для записи данных на машиночитаемый носитель, такое как, например, накопитель на оптических дисках, накопитель на магнитных носителях (например, накопитель на гибких дисках), порт универсальной последовательной шины (USB), сетевой интерфейс или другой интерфейс вывода. Интерфейс 32 вывода выводит видеофайл в машиночитаемый носитель, такой как, например, передаваемый сигнал, магнитный носитель, оптический носитель, запоминающее устройство, флэш-накопитель или другой машиночитаемый носитель.

Сетевой интерфейс 54 может принимать NAL-единицу или единицу доступа через сеть 74 и обеспечивать NAL-единицу или единицу доступа в модуль 50 декапсуляции через модуль 52 извлечения. Модуль 50 декапсуляции может декапсулировать элементы видеофайла на составляющие PES-потoki, депакетизировать PES-потoki, чтобы извлекать кодированные данные, и отправлять кодированные данные в аудиодекодер 46 или в видеодекодер 48, в зависимости от того, являются кодированные данные частью аудио-или видеопотока, например, как указано посредством заголовков PES-пакетов потока. Аудиодекодер 46 декодирует кодированные аудиоданные и отправляет декодированные аудиоданные в аудиовывод 42, тогда как видеодекодер 48 декодирует кодированные видеоданные и отправляет декодированные видеоданные, которые могут включать в себя множество видов потока, в видеовывод 44.

В соответствии с технологиями этого раскрытия сущности, любые из устройства 20 подготовки контента, серверного устройства 60 и/или клиентского устройства 40 могут быть выполнены с возможностью осуществлять различные способы для задания, передачи в сигналах и/или обработки мультимедийных данных согласно новому DASH-профилю (например, усовершенствованному профилю передач *вживую*). Аналогично, любые из этих устройств могут быть выполнены с возможностью обрабатывать новые типы мультимедийных сегментов, которые могут обеспечивать потоковую передачу видео с задержкой, включающую в себя уменьшенное время переключения каналов при широкоэмитательной передаче и многоадресной передаче, при одновременном обеспечении структур по стандарту высокоэффективного кодирования видео. В общем, поясняются следующие аспекты, которые могут выполняться отдельно или в любой комбинации:

Задание различных типов мультимедийных сегментов и их структур.

Анализ текущих атрибутов.

Вопросы, связанные с решениями.

Передача MPD-сигналов.

Передача в сигналах типа в сегменте.

Передача в сигналах типа в MPD.

Обеспечение адаптивных наборов для различных вариантов использования.

В некоторых примерах, устройство 20 подготовки контента, серверное устройство 60 и клиентское устройство 40 могут быть выполнены с возможностью использовать мультимедийные сегменты, соответствующие любому из следующих форматов: формат мультимедийных сегментов единиц доставки, формат мультимедийных сегментов с произвольным доступом, формат сегментов без перекрытия и/или формат мультимедийных сегментов с перекрытием. Ниже подробнее описываются эти форматы.

Мультимедийный сегмент, соответствующий формату мультимедийных сегментов единиц достав-

ки, может задаваться следующим образом:

Каждый мультимедийный сегмент должен содержать один или более целых автономных кинофрагментов. Полный автономный кинофрагмент представляет собой поле кинофрагментов (moof) и поле мультимедийных данных (mdat), которые содержат все мультимедийные выборки, которые не используют внешние ссылки на данные, к которым обращаются посредством серий дорожек в поле кинофрагментов.

Каждое поле moof должно содержать, по меньшей мере, один фрагмент дорожки.

Поля moof не должны использовать внешние ссылки на данные, флаг default-base-is-moof должен задаваться, и смещение данных должно использоваться, т.е. base-data-offset-present не должен использоваться. Эта комбинация настроек может упоминаться как относительная адресация кинофрагментов для мультимедийных данных.

Каждый мультимедийный сегмент может переносить "dums" в поле типов сегментов (styp) в качестве совместимого бренда. Требования по соответствию этого бренда могут быть такими, как задано в этом раскрытии сущности.

Мультимедийный сегмент, соответствующий формату мультимедийных сегментов с произвольным доступом, задается следующим образом:

Мультимедийный сегмент должен соответствовать формату мультимедийных сегментов единиц доставки, как указано выше.

Первая единица доступа в каждом кинофрагменте в мультимедийном сегменте с произвольным доступом должна соответствовать I_{SAU} SAP типа 1, 2 или 3 (например, включать в себя IDR-, CRA- или VLA-изображение).

Мультимедийный сегмент должен переносить достаточную информацию, чтобы осуществлять доступ к мультимедиа в потоке, например, все необходимое шифрование в комбинации с сегментом инициализации, при наличии.

Каждое поле traf (поле фрагментов дорожек) должно содержать поле tfdt (поле времени декодирования фрагментов дорожек).

Каждый мультимедийный сегмент может переносить "rams" в поле типов сегментов (styp) в качестве совместимого бренда. Требования по соответствию этого бренда задаются в этом подразделе.

Каждый мультимедийный сегмент может содержать одно или более полей sidx. Если присутствует, первое поле sidx должно быть размещено до любого поля moof, и первое поле индексов сегментов должно документировать весь сегмент.

Мультимедийный сегмент, соответствующий формату сегментов без перекрытия, может задаваться следующим образом:

Мультимедийный сегмент должен соответствовать формату мультимедийных сегментов единиц доставки, как указано выше.

Сегмент должен удовлетворять свойству отсутствия перекрытия, как задано в 4.5.3 из ISO/IEC 23009-1, в том смысле, что сегмент и его предыдущий сегмент удовлетворяют свойству отсутствия перекрытия.

Мультимедийный сегмент, соответствующий формату мультимедийных сегментов с переключением, может задаваться следующим образом:

Мультимедийный сегмент должен соответствовать формату мультимедийных сегментов с произвольным доступом, как указано выше.

Первая выборка в первом кинофрагменте в мультимедийном сегменте с переключением должна соответствовать I_{SAU} SAP типа 1 или 2 (например, IDR-изображению).

Каждый мультимедийный сегмент может переносить "swms" в поле типов сегментов (styp) в качестве совместимого бренда. Требования по соответствию этого бренда задаются в этом подразделе.

Сегменты различных форматов могут выполнять различные функции. Например, мультимедийные сегменты единиц доставки, в общем, выполняют функцию доставки мультимедийных данных. В качестве других данных, мультимедийные сегменты с произвольным доступом выполняют функцию обеспечения точек произвольного доступа (включающих в себя данные инициализации) в представление, включающее в себя мультимедийные сегменты с произвольным доступом. Сегменты без перекрытия могут выполнять функцию указания совмещения сегментов между представлениями, что позволяет обеспечить простое переключение представлений. Мультимедийные сегменты с переключением обеспечивают функцию обеспечения возможности переключения представлений без включения дополнительных данных инициализации, которые требуются для мультимедийного сегмента с произвольным доступом.

Кроме того, устройство 20 подготовки контента, серверное устройство 60 и клиентское устройство 40 могут быть выполнены с возможностью обрабатывать данные, представляющие форматы, поясненные выше, и/или другие данные согласно технологиям этого раскрытия сущности, например, в файле 66 манифеста (к примеру, MPD). Следующие признаки могут передаваться в сигналах в файле 66 манифеста, отдельно или в любой комбинации:

Тип каждого мультимедийного сегмента в представлении, явно передаваемого в сигналах или передаваемого в сигналах через шаблон.

Возможность иметь различные размеры сегментов в одном адаптивном наборе, но при этом иметь совмещенные точки переключения, т.е. мультимедийные сегменты с переключением, начинающиеся одновременно.

Последствия для вычисления `minBufferTime` и полосы пропускания (должно начинаться в точке произвольного доступа).

Для каждого из представлений 68 и, возможно, на установленном по умолчанию уровне адаптивных наборов, следующее может передаваться в сигналах в файле 66 манифеста:

Шаблон в представлении:

Каждый сегмент имеет мультимедийный сегмент единиц доставки типа, каждый N-й сегмент представляет собой мультимедийный сегмент с произвольным доступом, каждый M-й представляет собой сегмент с переключением, где $M \geq N$. Некоторые сокращения и настройки по умолчанию могут быть выполнимыми.

Это может передаваться в сигналах с новым атрибутом: `rams-frequency` и `swms-frequency`.

Другие шаблоны сокращения, которые обеспечивают возможность выражения шаблона без обновления MPD.

Шаблон на временной шкале сегментов.

Добавление необязательного поля типа на временной шкале сегментов для каждого элемента.

Тип сегмента.

Поле типа также может выражать шаблон в качестве вышеуказанного шаблона.

Обеспечение возможности передавать в сигналах нерегулярности с обновлениями S-элемента на временной шкале сегментов.

Явное

Добавление поля, которое обеспечивает возможность передачи в сигналах шаблонов сегментов в явном списке, возможно с чередованием с некоторыми шаблонами.

Это также может включать в себя передачу в сигналах длительности сегмента.

Может иметь место то, что представления в общем адаптивном наборе имеют различные длительности сегментов. Тем не менее, проблема для переключения заключается в том, что точки переключения между представлениями должны совмещаться, с тем чтобы обеспечивать прозрачное переключение. Позиция точек переключения может передаваться в сигналах, как пояснено выше. Следующая передача сигналов также может рассматриваться:

Все представления имеют точки переключения в идентичной позиции, и они совмещаются. Это может передаваться в сигналах с помощью одного флага.

Когда точка переключения передается в сигналах в конкретное время (в этом случае, в MPD-время, которое может быть сложным), то она совмещается со всеми другими точками переключения в представлении. Это также может передаваться в сигналах с помощью одного флага, и может использоваться идентичный флаг, как пояснено выше.

В некоторых примерах, даже в случае, если нет последующего мультимедийного сегмента с переключением, по-прежнему отсутствует перекрытие, так что клиентское устройство 40 может переключаться с точки без перекрытия на мультимедийный сегмент с переключением.

Другая более явная передача сигналов точек переключения дополнительно может передаваться в сигналах в файле 66 манифеста.

Как отмечено выше, в некоторых примерах, устройство 20 подготовки контента, серверное устройство 60 и/или клиентское устройство 40 могут быть выполнены с возможностью использовать усовершенствованный профиль передач вживую DASH. Усовершенствованный профиль передач вживую может включать в себя все признаки и типы сегментов, заданные выше. Усовершенствованный профиль передач вживую может идентифицироваться посредством универсального имени ресурса (URN): `"urn:mpeg:dash:profile:advanced-live:2015"`.

В некоторых примерах, если усовершенствованный профиль передач вживую используется в адаптивном наборе:

Каждый мультимедийный сегмент с переключением должен переносить "swms" в поле типов сегментов (`styp`) в качестве совместимого бренда.

Каждый мультимедийный сегмент с произвольным доступом, который не переносит "swms", должен переносить "rams" в поле типов сегментов (`styp`) в качестве совместимого бренда.

Это раскрытие сущности распознает следующие проблемы и ограничения для традиционной передачи сигналов для MPD-атрибутов:

1. Передача в сигналах времени доступности сегментов:

@duration или временная шкала сегментов:

Предложение состоит в том, чтобы выполнять упрощение в новом профиле и использовать для этой цели только временную шкалу сегментов, поскольку она представляет собой расширенный набор @duration.

Тем не менее, временная шкала сегментов является более сложной, поскольку она разрешает исключения.

Также необходимо четко прояснять, является время на временной шкале сегментов точной длительностью сегмента (обеспечивает меньшую гибкость в авторской разработке контента) или длительностью без уходов, и передает в сигналах только времена доступности сегментов.

Важно отметить, что посредством надлежащего применения @timescale, эта проблема может решаться.

2. Передача в сигналах переключения из свойства, т.е. без перекрытия

Посредством задания совмещения сегментов как истинного в адаптивном наборе.

Проблема заключается в том, что это означает то, что каждый сегмент должен иметь идентичную длительность.

Отсутствие перекрытия должно выражаться с большей степенью детализации.

3. Передача в сигналах произвольного доступа

Начала с SAP задаются равными 1, 2 или 3:

Проблема заключается в том, что это объявляется не очень явно.

Также должны задаваться другие требования, см. расширенное определение сегмента с произвольным доступом.

4. Передача в сигналах точки переключения.

Начала с SAP задаются равными 1 или 2:

Проблема заключается в том, что это объявляется не очень явно.

Может применяться другой тип переключения, но это требует дополнительных идей. Некоторая гибкость должна добавляться.

5. Передача в сигналах URL-адреса сегмента

Шаблон на основе номеров.

Проблема заключается в том, что по существу предполагается, что каждый сегмент имеет идентичный номер в каждом представлении в каждом адаптивном наборе. Следует отметить, что это не требование, а вероятное допущение в реализациях. В случае такого изменения, чтобы иметь различные размеры сегментов в одном адаптивном наборе, больше нет соответствия по нумерации.

Для простоты на данный момент, номера не используются.

Временной шаблон

Проблема заключается в том, что по существу предполагается, что каждый сегмент имеет идентичное время в каждом представлении в каждом адаптивном наборе. Следует отметить, что это не требование, а вероятное допущение в реализациях.

Тем не менее, также следует отметить, что это может выражаться на общей временной шкале. Кроме того, временная шкала является более подходящей, чем нумерация, для того чтобы выразить взаимосвязь между различными представлениями.

Список сегментов

Проблема заключается в том, что здесь позиция списка совмещает сегменты, и может возникать случай, в котором именованное является произвольным. Клиент должен поддерживать точное преобразование списка и порядок каждого представления в адаптивном наборе.

Это раскрытие сущности технологии для назначения различных фрагментов по мере необходимости. Серверное устройство 60 и клиентское устройство 40 могут быть сконфигурированы согласно следующему подходу в некоторых примерах:

Длительность/временная шкала сегментов назначается единице доставки, поскольку она выражает время, когда сегмент доступен в сервере.

Время не может быть точным с точки зрения времени в мультимедиа, а используется для того, чтобы вычислять начальное время доступности сегментов.

Эта временная синхронизация может отличаться для различного представления в одном адаптивном наборе. Например, могут быть предусмотрены представления, которые доступны с большим числом единиц доставки, чем другие. См. пояснение варианта использования.

Необходимы четкие инструкции касательно того, как вычислять начальное время доступности сегментов на основе вышеуказанных сигналов. Существующая модель является эффективной, но при практическом использовании существующая модель должна, безусловно, использоваться надлежащим образом, если существующая модель для вычисления начального времени доступности сегментов должна использоваться в соответствии с технологиями этого раскрытия сущности.

Это включает в себя то, что время доступности сегментов может регулироваться для определенных представлений или базовых URL-адресов посредством смещения времени доступности.

Другой важный вопрос, который следует прояснить, заключается в том, как нерегулярная длительность сегментов затрагивает начальное время доступности и передачу сигналов. Обычно, сегменты должны иметь идентичный размер.

Произвольный доступ может отличаться в различных представлениях.

Необходимо прояснить то, находится произвольный доступ по времени только в начале сегмента, либо он также может находиться в середине сегмента.

Согласно 4.2.2, он в настоящее время находится в начале сегмента, но это может приводить к нере-

гулярным размерам сегментов, если точки произвольного доступа являются нерегулярно размещенными. Это снова затрагивает время задержки, поскольку доступность сегментов является менее прогнозируемой.

Тем не менее, в качестве рабочего допущения, которое должна поддерживать модель 4.2.2, то, что произвольный доступ находится в начале сегмента.

Произвольный доступ может передаваться в сигналах в двух областях, во времени или в нумерации сегментов.

Чтобы достигать общего инструментального средства, может использоваться временной подход.

По меньшей мере, два подхода к переключению пояснены в базовых экспериментах:

Переключение потоков битов:

DASH-клиент не знает внутренние структуры представлений. Он знает только то, где он может сращивать представления, и подает означенное в декодер мультимедиа в качестве одного потока битов. Кодер удостоверяется в том, что представления кодируются таким образом, что это свойство удовлетворяется на уровне инкапсуляции и мультимедийных потоков.

Это по существу должно разрешать клиенту создавать последовательность/поток битов следующим образом:

Сегмент инициализации для адаптивного набора

Мультимедийный сегмент 1 представления 1

Мультимедийный сегмент X представления 1

...

Мультимедийный сегмент X+1 представления 2

...

Переключение обеспечивается посредством конкретных свойств в мультимедиа. Именно это осуществляется в DASH. Созданы некоторые правила относительно того, как переключение может выполняться на уровне воспроизведения файлов. Основное правило состоит в том, что известно то, что если совмещение сегментов задается как истинное, начало с SAP равно 1 или 2, следующая последовательность обеспечивает прозрачное переключение:

Сегмент инициализации представления 1

Мультимедийный сегмент 1 представления 1

...

Мультимедийный сегмент X представления 1

Сегмент инициализации представления 2

Мультимедийный сегмент X+1 представления 2

...

Переключение в открытой GOP или другие аспекты, которые требуют более подробного понимания обработки мультимедиа.

Расширения и ограничения могут применяться к файлу бб манифеста (например, MPD) на основе вышеприведенного пояснения, при этом расширения и ограничения могут применяться к новым инструментальным средствам. Например, следующие расширения могут применяться, отдельно или в любой комбинации:

Добавление нового атрибута `@randomAccessPeriod` (или любого другого средства для того, чтобы выражать период произвольного доступа), который выражается на шкале `@timescale` на уровне представления. Любой сегмент, для которого `$Time$` падает до целого кратного произведения `@timescale` и `@randomAccessPeriod`, представляет собой сегмент с произвольным доступом, т.е. он разрешает осуществлять доступ к адаптивному набору этому представлению.

Произвольный доступ может квалифицироваться дополнительно, например, то, какой SAP-тип доступен в какой период, т.е. SAP-тип 1, 2 или 3. Следует отметить, что 3 означает то, что опытный SAP-тип также может быть равным 1 или 2.

Добавление нового элемента: мультимедийного сегмента с переключением (или любого другого средства для того, чтобы выражать переключение) с двумя атрибутами на уровне адаптивных наборов (могут присутствовать один или более):

`@period`, выражаемый на шкале `@timescale`. Любая временная позиция, для которой `$Time$` падает до целого кратного произведения `@timescale` и обеспечивает возможность переключения, т.е. он разрешает переключаться на это представление.

`@type`, выражающий тип переключения, активируется. Задаются, по меньшей мере, два типа, а именно, переключение потоков битов и переключение мультимедийных уровней. Могут задаваться другие типы, такие как переключение открытых GOP.

Другой способ выразить такое переключение состоит в том, чтобы использовать тип дескриптора, при этом дескриптор выражает тип переключения и значение частоты переключения.

На временной шкале сегментов и S-элементе, обеспечение дополнительного атрибута `@reset`, который по умолчанию задается равным лжи. Сброс означает, что периодичность периода произвольного доступа и периода переключения сбрасывается в этой точке. Это обеспечивает возможность того, что

IDR добавляются, и временная шкала сегментов по существу сбрасывается в более произвольные времена.

Вышеуказанный сценарий не обязательно поддерживает вариант использования, в котором шаблоны сегментов обеспечивают доступности сегментов, поясненные выше. Чтобы также разрешать этот вариант использования, может добавляться следующее расширение:

Добавление нового элемента переключения (или любого другого средства или элемента для того, чтобы выражать переключение) с двумя атрибутами на уровне представления (могут присутствовать один или более):

`@period`, выражаемый на шкале `@timescale`. Любая временная позиция, для которой `$Time$` падает до целого кратного произведения `@timescale` и обеспечивает возможность переключения, т.е. он разрешает переключаться на это представление.

`@type`, выражающий тип переключения, активируется. Задаются, по меньшей мере, два типа, а именно, переключение потоков битов и переключение мультимедийных уровней. Могут задаваться другие типы, такие как переключение открытых GOP.

Следующие ограничения предложены для применения для усовершенствованного профиля передач вживую, чтобы обеспечивать усовершенствованные варианты использования:

Использование одного `@timescale` для всех представлений в одном адаптивном наборе.

Использование временной шкалы сегментов для передачи в сигналах длительностей сегментов (для простоты).

Использование только `$Time$` для передачи в сигналах URL-адреса (теперь для простоты).

Временная синхронизация по длительности сегмента является точной (рабочее допущение, необходимо понимать последствия).

Точность длительности сегмента может управляться посредством `@timescale` при использовании (примечание), например, если временная шкала составляет только 1/5 от фактической частоты дискретизации, имеется некоторая гибкость по точной частоте дискретизации.

Временная шкала сегментов задается в расчете на представление, чтобы обеспечивать возможность различных длительностей сегментов в различных представлениях. Тем не менее, это может быть установлено по умолчанию на уровне адаптивных наборов.

Временная шкала сегментов может использовать открытый `@r(-1)` или закрытый `@r(>=0)`.

Совмещение сегментов и начало с SAP могут использоваться для обратно совместимых развертываний, но, в общем, не должны использоваться. Передача сигналов должна всегда обеспечиваться посредством `@randomAccessPeriod` и элемента с переключением.

Необходимо обеспечивать то что, если адаптивный набор содержит одно, а не большее число представлений, то логика переключения обеспечивается для представления на уровне адаптивных наборов.

Хотя главным образом описаны относительно DASH, технологии этого раскрытия сущности также могут использоваться для других форматов мультимедиа, таких как MPEG-2 TS (транспортный поток) или WebM.

Таким образом, клиентское устройство 40 представляет пример устройства для извлечения мультимедийных данных, содержащего один или более процессоров, выполненных с возможностью извлекать мультимедийный сегмент, соответствующий, по меньшей мере, одному из формата мультимедийных сегментов единиц доставки, формата мультимедийных сегментов с произвольным доступом, формата сегментов без перекрытия или формата мультимедийных сегментов с переключением, и обрабатывать мультимедийный сегмент, по меньшей мере, частично на основе того, соответствует мультимедийный сегмент формату мультимедийных сегментов единиц доставки, формату мультимедийных сегментов с произвольным доступом, формату сегментов без перекрытия или формату мультимедийных сегментов с переключением.

Клиентское устройство 40 также представляет пример устройства для извлечения мультимедийных данных, содержащего один или более процессоров, выполненных с возможностью принимать файл манифеста, включающий в себя данные, указывающие шаблон для мультимедийных сегментов различных типов в представлении, и извлекать один или более мультимедийных сегментов, по меньшей мере, частично на основе шаблона.

Кроме того, клиентское устройство 40 представляет пример устройства для извлечения мультимедийных данных, содержащего один или более процессоров, выполненных с возможностью определять, из файла манифеста, множество типов сегментов, включенных в представление мультимедийного контента, причем одна или более функций обеспечиваются посредством каждого из типов сегментов и позиций сегментов, соответствующих каждому из типов сегментов в представлении, при этом, по меньшей мере, один из типов сегментов обеспечивает точку, в которой можно начинать извлечение данных из представления, определять, из файла манифеста, сегмент представления, соответствующий типу, который обеспечивает точку, в которой можно начинать извлечение данных из представления, и извлекать определенный сегмент из представления.

Аналогично, серверное устройство 60 и устройство 20 подготовки контента представляют примеры устройства для отправки мультимедийных данных, причем устройство содержит один или более процес-

соров, выполненных с возможностью формировать мультимедийный сегмент, соответствующий, по меньшей мере, одному из формата мультимедийных сегментов единиц доставки, формата мультимедийных сегментов с произвольным доступом, формата сегментов без перекрытия или формата мультимедийных сегментов с переключением, и отправлять мультимедийный сегмент в клиентское устройство.

Серверное устройство 60 и устройство 20 подготовки контента также представляют примеры устройства для отправки мультимедийных данных, причем устройство содержит один или более процессоров, выполненных с возможностью отправлять файл манифеста, включающий в себя данные, указывающие шаблон для мультимедийных сегментов различных типов в представлении, в клиентское устройство и отправлять, в ответ на один или более запросов, один или более мультимедийных сегментов, по меньшей мере, частично на основе шаблона в клиентское устройство.

Серверное устройство 60 и устройство 20 подготовки контента также представляют примеры устройства для передачи в сигналах мультимедийной информации, причем устройство включает в себя один или более процессоров, выполненных с возможностью составлять файл манифеста, указывающий множество типов сегментов, включенных в представление мультимедийного контента, причем одна или более функций обеспечиваются посредством каждого из типов сегментов, причем позиции сегментов соответствуют каждому из типов сегментов в представлении, при этом, по меньшей мере, один из типов сегментов обеспечивает точку, в которой можно начинать извлечение данных из представления, и сегмент представления, соответствующего типу, который обеспечивает точку, в которой можно начинать извлечение данных из представления, отправлять файл манифеста в клиентское устройство, и в ответ на запрос из клиентского устройства на предмет сегмента, соответствующего типу, который обеспечивает точку, в которой можно начинать извлечение данных из представления, отправлять сегмент, который обеспечивает точку, в которой можно начинать извлечение данных из представления, в клиентское устройство.

Фиг. 5А является концептуальной схемой, иллюстрирующей элементы примерного мультимедийного контента 102. Мультимедийный контент 102 может соответствовать мультимедийному контенту 64 (Фиг. 4) или другому мультимедийному контенту, сохраненному на носителе 62 хранения данных. В примере по фиг. 5А, мультимедийный контент 102 включает в себя описание 104 мультимедийного представления (MPD) и множество представлений 110А-110N (представлений 110). Представление 110А включает в себя необязательные данные 112 заголовка и сегменты 114А-114N (сегменты 114), тогда как представление 110N включает в себя необязательные данные 122 заголовка и сегменты 124А-124N (сегменты 124). Буква N используется для того, чтобы обозначать последний кинофрагмент в каждом из представлений 110 для удобства. В некоторых примерах, могут быть предусмотрены различные числа кинофрагментов между представлениями 110.

MPD 104 может содержать структуру данных, отдельную от представлений 110. MPD 104 может соответствовать файлу 66 манифеста по фиг. 4. Аналогично, представления 110 могут соответствовать представлениям 68 по фиг. 4. В общем, MPD 104 может включать в себя данные, которые, в общем, описывают характеристики представлений 110, такие как характеристики кодирования и рендеринга, адаптивные наборы, профиль, которому соответствует MPD 104, информация типов текста, информация ракурсов камеры, рейтинговая информация, информация режимов трюков (например, информация, указывающая представления, которые включают в себя временные подпоследовательности), и/или информация для извлечения удаленных периодов (например, для вставки адресных рекламных объявлений в мультимедийный контент в ходе воспроизведения).

Данные 112 заголовка, когда присутствуют, могут описывать характеристики сегментов 114, например, временные местоположения точек произвольного доступа (RAP, также называемых в качестве точек доступа к потоку (SAP)), то, какие из сегментов 114 включают в себя точки произвольного доступа, байтовые смещения в точках произвольного доступа в сегментах 114, универсальные указатели ресурса (URL-адреса) сегментов 114 либо другие аспекты сегментов 114. Данные 122 заголовка, когда присутствуют, могут описывать аналогичные характеристики для сегментов 124. Дополнительно или альтернативно, такие характеристики могут быть полностью включены в MPD 104.

Сегменты 114, 124 включают в себя одну или более кодированных видеовыборок, каждая из которых может включать в себя кадры или серии последовательных макроблоков видеоданных. Каждая из кодированных видеовыборок сегментов 114 может иметь аналогичные характеристики, например, высоту, ширину и требования по полосе пропускания. Такие характеристики могут описываться посредством данных MPD 104, хотя такие данные не проиллюстрированы в примере по фиг. 5А. MPD 104 может включать в себя характеристики, как описано посредством технических требований 3GPP с добавлением любой из передаваемой в сигналах информации, описанной в этом раскрытии сущности.

Каждый из сегментов 114, 124 может быть ассоциирован с уникальным универсальным указателем ресурса (URL-адресом). Таким образом, каждый из сегментов 114, 124 может быть независимо извлекаемым с использованием сетевого протокола потоковой передачи, такого как DASH. Таким образом, устройство назначения, к примеру, клиентское устройство 40, может использовать HTTP GET-запрос, чтобы извлекать сегменты 114 или 124. В некоторых примерах, клиентское устройство 40 может использовать частичные HTTP GET-запросы, чтобы извлекать конкретные байтовые диапазоны сегментов 114 или 124.

Фиг. 5В является концептуальной схемой, иллюстрирующей примерный контент описания 104 мультимедийного представления (MPD) в соответствии с технологиями этого раскрытия сущности. В общем, в числе других данных, передаваемых в сигналах в MPD 104, в примере по фиг. 5В, MPD 104 включает в себя информацию 130 периода, информацию 132 адаптивного набора и информацию 134А-134N представления (информацию 134 представления). Хотя только один набор информации 132 адаптивного набора показан в этом примере, следует понимать, что, в общем, может быть включено множество наборов информации адаптивного набора. Аналогично, хотя показан только один набор информации 130 периода, следует понимать, что, в общем, может быть включено множество наборов информации периода.

В соответствии с технологиями этого раскрытия сущности, информация 134А представления включает в себя информацию 136А типов сегментов, информацию 138А функций сегментов и местоположения 140А сегментов. Аналогично, информация 134N представления включает в себя информацию 136N типов сегментов, информацию функций сегментов 138N и местоположения 140N сегментов. В общем, информация 136А, 136N типов сегментов описывает различные типы сегментов, включенных в представление, соответствующие информации 134А, 134N представления, соответственно. Например, типы 136А, 136N сегментов могут включать в себя любое из типа (или формата) мультимедийных сегментов единиц доставки, типа (или формата) мультимедийных сегментов с произвольным доступом, типа (или формата) сегментов без перекрытия и типа (или формата) мультимедийных сегментов с переключением.

Информация 138А, 138N функций сегментов, в общем, описывает функции, выполняемые посредством различных типов сегментов. Например, информация 138А, 138N функций сегментов может описывать функции, выполняемые посредством любого из типа (или формата) мультимедийных сегментов единиц доставки, типа (или формата) мультимедийных сегментов с произвольным доступом, типа (или формата) сегментов без перекрытия и типа (или формата) мультимедийных сегментов с переключением, при условии, что такие типы/форматы присутствуют в соответствующей информации 136А, 136N типов сегментов. Информация 138А, 138N функций сегментов может указывать то, что тип мультимедийных сегментов единиц доставки используется для того, чтобы, в общем, переносить мультимедийные данные, тип мультимедийных сегментов с произвольным доступом используется для того, чтобы обеспечивать точку произвольного доступа (которая включает в себя информацию инициализации), тип сегментов без перекрытия указывает то, что такие сегменты не перекрывают другие сегменты идентичного представления или других представлений, и тип мультимедийных сегментов с переключением обеспечивает возможность переключения между представлениями в адаптивном наборе.

Информация 140А, 140N местоположений сегментов, в общем, может передавать в сигналах местоположения (или позиции) сегментов различных типов в соответствующих представлениях. Например, информация 140А, 140N местоположений сегментов может передавать в сигналах частоты, с которыми сегменты каждого из типа мультимедийных сегментов единиц доставки, типа мультимедийных сегментов с произвольным доступом, типа сегментов без перекрытия и/или типа мультимедийных сегментов с переключением возникают в соответствующих представлениях. Информация 140А, 140N местоположений сегментов может указывать эту информацию в форме шаблона (например, каждый N-й сегмент представляет собой сегмент типа X). Дополнительно или альтернативно, информация 140А, 140N местоположений сегментов может явно перечислять местоположения отдельных сегментов.

Фиг. 6 является блок-схемой, иллюстрирующей элементы примерного видеофайла 150, которые могут соответствовать сегменту представления, такому как один из сегментов 114, 124 по фиг. 5А. Каждый из сегментов 114, 124 может включать в себя данные, которые практически соответствуют компоновке данных, проиллюстрированных в примере по фиг. 6. Можно сказать, что видеофайл 150 инкапсулирует сегмент. Как описано выше, видеофайлы в соответствии с базовым форматом мультимедийных файлов ISO и его расширениями сохраняют данные в последовательности объектов, называемых в качестве "полей". В примере по фиг. 6, видеофайл 150 включает в себя поле 152 типов файлов (FTYP), кинополе 154 (MOOV), поля 162 индексов сегментов (sidx), поля 164 кинофрагментов (MOOF) и поле 166 произвольного доступа к кинофрагментам (MFRA). Хотя фиг. 6 представляет пример видеофайла, следует понимать, что другие мультимедийные файлы могут включать в себя другие типы мультимедийных данных (например, аудиоданные, синхронизированные по времени текстовые данные и т.п.), которые структурированы аналогично данным видеофайла 150, в соответствии с базовым форматом мультимедийных файлов ISO и его расширениями.

Поле 152 типов файлов (FTYP), в общем, описывает тип файла для видеофайла 150. Поле 152 типов файлов может включать в себя данные, которые идентифицируют спецификацию, которая описывает наилучшее использование для видеофайла 150. Поле 152 типов файлов альтернативно может быть размещено перед полем 154 MOOV, полями 164 кинофрагментов и/или полем 166 MFRA.

В некоторых примерах, сегмент, такой как видеофайл 150, может включать в себя поле MPD-обновления (не показано) до поля 152 Ftup. Поле MPD-обновления может включать в себя информацию, указывающую то, что MPD, соответствующее представлению, включающему в себя видеофайл 150, должно обновляться, вместе с информацией для обновления MPD. Например, поле MPD-обновления может обеспечивать URI или URL-адрес для ресурса, который должен использоваться для того, чтобы

обновлять MPD. В качестве другого примера, поле MPD-обновления может включать в себя данные для обновления MPD. В некоторых примерах, поле MPD-обновления может идти сразу после поля типов сегментов (STYP) (не показано) видеофайла 150, при этом поле STYP может задавать тип сегмента для видеофайла 150. фиг. 7, подробнее поясненный ниже, обеспечивает дополнительную информацию относительно поля MPD-обновления.

Поле 154 MOOV, в примере по фиг. 6, включает в себя поле 156 кинозаголовков (MVHD), поле 158 дорожек (TRAK) и одно или более полей 160 кинорасширений (MVEX). В общем, поле 156 MVHD может описывать общие характеристики видеофайла 150. Например, поле 156 MVHD может включать в себя данные, которые описывают то, когда видеофайл 150 первоначально создан, когда видеофайл 150 в последний раз модифицирован, временную шкалу для видеофайла 150, длительность воспроизведения для видеофайла 150 либо другие данные, которые, в общем, описывают видеофайл 150.

Поле 158 TRAK может включать в себя данные для дорожки видеофайла 150. Поле 158 TRAK может включать в себя поле (TKHD) заголовков дорожек, которое описывает характеристики дорожки, соответствующей полю 158 TRAK. В некоторых примерах, поле 158 TRAK может включать в себя кодированные видеоизображения, тогда как в других примерах, кодированные видеоизображения дорожки могут быть включены в кинофрагменты 164, к которым можно обратиться посредством данных поля 158 TRAK и/или полей 162 SIDX.

В некоторых примерах, видеофайл 150 может включать в себя более одной дорожки. Соответственно, поле 154 MOOV может включать в себя число полей TRAK, равное числу дорожек в видеофайле 150. Поле 158 TRAK может описывать характеристики соответствующей дорожки видеофайла 150. Например, поле 158 TRAK может описывать временную и/или пространственную информацию для соответствующей дорожки. Поле TRAK, аналогичное полю 158 TRAK поля 154 MOOV, может описывать характеристики дорожки наборов параметров, когда модуль 30 инкапсуляции (Фиг. 4) включает дорожку наборов параметров в видеофайл, к примеру, в видеофайл 150. Модуль 30 инкапсуляции может передавать в сигналах присутствие SEI-сообщений уровня последовательности в дорожке наборов параметров в поле TRAK, описывающем дорожку наборов параметров.

Поля 160 MVEX могут описывать характеристики соответствующих кинофрагментов 164, например, передавать в сигналах то, что этот видеофайл 150 включает в себя кинофрагменты 164, в дополнение к видеоданным, включенным в поле 154 MOOV, если таковые имеются. В контексте потоковой передачи видеоданных, кодированные видеоизображения могут быть включены в кинофрагменты 164, а не в поле 154 MOOV. Соответственно, все кодированные видеовыборки могут быть включены в кинофрагменты 164, а не в поле 154 MOOV.

Поле 154 MOOV может включать в себя число полей 160 MVEX, равное числу кинофрагментов 164 в видеофайле 150. Каждое из полей 160 MVEX может описывать характеристики соответствующего одного из кинофрагментов 164. Например, каждое поле MVEX может включать в себя поле заголовков кинорасширений (MHND), которое описывает временную длительность для соответствующего одного из кинофрагментов 164.

Как отмечено выше, модуль 30 инкапсуляции может сохранять набор данных последовательности в видеовыборке, которая не включает в себя фактические кодированные видеоданные. Видеовыборка, в общем, может соответствовать единице доступа, которая является представлением кодированного изображения в конкретный момент времени. В контексте AVC, кодированное изображение включает в себя одну или более VCL NAL-единиц, которые содержат информацию для того, чтобы составлять все пиксели единицы доступа, и других ассоциированных не-VCL NAL-единиц, таких как SEI-сообщения. Соответственно, модуль 30 инкапсуляции может включать в себя набор данных последовательности, который может включать в себя SEI-сообщения уровня последовательности, в одном из кинофрагментов 164. Модуль 30 инкапсуляции дополнительно может передавать в сигналах присутствие набора данных последовательности и/или SEI-сообщения уровня последовательности как присутствующие в одном из кинофрагментов 164 в пределах одного из полей 160 MVEX, соответствующих одному из кинофрагментов 164.

Поля 162 SIDX являются необязательными элементами видеофайла 150. Иными словами, видеофайлы, соответствующие 3GPP-формату файлов или другим таким форматам файлов, не обязательно включают в себя поля 162 SIDX. В соответствии с примером 3GPP-формата файлов, поле SIDX может использоваться для того, чтобы идентифицировать субсегмент сегмента (например, сегмента, содержащегося в видеофайле 150). 3GPP-формат файлов задает субсегмент в качестве "автономного набора из одного или более последовательных полей кинофрагментов с соответствующим полем(ями) мультимедийных данных, и поле мультимедийных данных, содержащее данные, к которым обращаются посредством поля кинофрагментов, должно следовать после этого поля кинофрагментов и предшествовать следующему полю кинофрагментов, содержащему информацию относительно идентичной дорожки". 3GPP-формат файлов также указывает то, что поле SIDX "содержит последовательность ссылок на субсегменты для (суб)сегмента, задокументированного посредством поля. Субсегменты, к которым обращаются, являются смежными во времени представления. Аналогично, байты, на которые ссылаются посредством поля индексов сегментов, всегда являются смежными в сегменте. Размер, к которому обращаются, обес-

печивает подсчет числа байтов в материале, к которому обращаются".

Поля 162 SIDX, в общем, обеспечивают информацию, представляющую один или более субсегментов сегмента, включенного в видеофайл 150. Например, эта информация может включать в себя времена воспроизведения, в которые субсегменты начинаются и/или завершаются, байтовые смещения для субсегментов, то, включают или нет субсегменты в себя (например, начинаются) точку доступа к потоку (SAP), тип для SAP (например, то, представляет собой SAP изображение на основе мгновенного обновления декодера (IDR), изображение на основе чистого произвольного доступа (CRA), изображение на основе доступа к нерабочей ссылке (BLA) и т.п.), позицию SAP (с точки зрения времени воспроизведения и/или байтового смещения) в субсегменте и т.п.

Кинофрагменты 164 могут включать в себя одно или более кодированных видеоизображений. В некоторых примерах, кинофрагменты 164 могут включать в себя одну или более групп изображений (GOP), каждая из которых может включать в себя определенное число кодированных видеоизображений, например, кадров или изображений. Помимо этого, как описано выше, кинофрагменты 164 могут включать в себя наборы данных последовательности в некоторых примерах. Каждый из кинофрагментов 164 может включать в себя поле заголовков кинофрагментов (MFHD, не показано на фиг. 6). Поле MFHD может описывать характеристики соответствующего кинофрагмента, такие как порядковый номер для кинофрагмента. Кинофрагменты 164 могут быть включены в порядке порядкового номера в видеофайл 150.

Поле 166 MFRA может описывать точки произвольного доступа в кинофрагментах 164 видеофайла 150. Это может помочь с выполнением режимов трюков, таких как выполнение поисков для конкретных временных местоположений (т.е. времен воспроизведения) в сегменте, инкапсулированном посредством видеофайла 150. Поле 166 MFRA, в общем, является необязательным и не должно включаться в видеофайлы, в некоторых примерах. Аналогично, клиентское устройство, к примеру, клиентское устройство 40, не обязательно должно ссылаться на поле 166 MFRA, чтобы корректно декодировать и отображать видеоданные видеофайла 150. Поле 166 MFRA может включать в себя число полей произвольного доступа к фрагментам дорожек (TFRA) (не показаны), равное числу дорожек видеофайла 150 или, в некоторых примерах, равное числу мультимедийных дорожек (например, дорожек, отличных от дорожек с подсказками) видеофайла 150.

В некоторых примерах, кинофрагменты 164 могут включать в себя одну или более точек доступа к потоку (SAP), таких как IDR-изображения. Аналогично, поле 166 MFRA может обеспечивать индикаторы относительно местоположений в видеофайле 150 SAP. Соответственно, временная подпоследовательность видеофайла 150 может формироваться из SAP видеофайла 150. Временная подпоследовательность также может включать в себя другие изображения, такие как P-кадры и/или B-кадры, которые зависят от SAP. Кадры и/или серии последовательных макроблоков временной подпоследовательности могут размещаться в сегментах таким образом, что кадры/серии последовательных макроблоков временной подпоследовательности, которые зависят от других кадров/серий последовательных макроблоков подпоследовательности, могут надлежащим образом декодироваться. Например, в иерархической компоновке данных, данные, используемые для прогнозирования для других данных, также могут быть включены во временную подпоследовательность.

Усовершенствованный профиль передач вживую является ожидаемым новым профилем, который акцентирует внимание на распространении услуг передач вживую. Ожидаемый профиль не обязательно считается обратно совместимым с расширенным общим профилем. Тем не менее, считается, что поставщик контента может формировать обратно совместимую версию контента, если считается важной. Чертежи, поясненные ниже, представляют различные варианты использования, в которых могут применяться технологии этого раскрытия сущности.

Фиг. 7 является концептуальной схемой, иллюстрирующей примерное предложение сегментов для варианта использования согласно технологиям этого раскрытия сущности. В частности, фиг. 7 иллюстрирует адаптивный набор 230, включающий в себя представление 232 и представление 234. Представление 232 включает в себя сегменты 236A-236E, которые включают в себя IDR-сегмент 236A и IDR-сегмент 236E, тогда как представление 234 включает в себя сегменты 238A-238A, которые включают в себя IDR-сегмент 238A и IDR-сегмент 238E.

Этот вариант использования включает в себя услуги потоковой передачи видео с низкой задержкой и переключение. Допустим, что сегмент составляет 0,5 секунды по длительности (с точки зрения времени воспроизведения), и частота кадров составляет 50 кадров в секунду (FPS). В этом примере и на основе технологий этого раскрытия сущности, установление и передача сигналов могут заключаться в следующем:

Каждый четвертый сегмент представляет собой сегмент с переключением/IDR (на основе мгновенного обновления декодера) Каждый сегмент представляет собой единицу доставки Передача сигналов может заключаться в следующем для адаптивного набора 230 согласно фиг. 7:

```

- AdaptationSet
-- @timescale=50
-- SegmentTimeline.S: @t=0, @d=25, @r=-1
-- @randomAccessPeriod=100
-- Switching: @period=100, @type="media"
--
SegmentTemplate@media="http://example.com/$RepresentationID$"/segment_
$Time$.mp4
- Representation: @id=232
- Representation: @id=234

```

Другой вариант использования согласно технологиям этого раскрытия сущности, включающий в себя услуги потоковой передачи видео с низкой задержкой и переключение, описывается относительно фиг. 1. фиг. 1 иллюстрирует предложение сегментов в случае этого варианта использования. Допустим, что короткий сегмент составляет 0,5 секунды по длительности, и частота кадров составляет 50 FPS. На основе технологий, поясненных выше, установление и передача сигналов для этого варианта использования могут заключаться в следующем:

Каждый сегмент представляет собой сегмент с произвольным доступом.

Сегменты в представлении широковещательной передачи в четыре раза превышают по размеру сегменты в представлении одноадресной передачи.

Сегмент в позиции перекрытия широковещательной передачи/одноадресной передачи представляет собой сегмент с переключением.

Передача сигналов может заключаться в следующем для адаптивного набора 230 согласно фиг. 7:

```

- AdaptationSet
-- @timescale=50
-- Switching: @period=100, @type="media"
--
SegmentTemplate@media="http://example.com/$RepresentationID$"/segment_
$Time$.mp4
- Representation: @id=1, @randomAccessPeriod=100
- SegmentTimeline.S: @t=0, @d=100, @r=-1
- Representation: @id=2, @randomAccessPeriod=25
- SegmentTimeline.S: @t=0, @d=25, @r=-1

```

Фиг. 8 является концептуальной схемой, иллюстрирующей вариант использования, включающий в себя быструю настройку с помощью масштабируемого HEVC (SHVC) в соответствии с технологиями этого раскрытия сущности. Пример по фиг. 8 иллюстрирует адаптивный набор 240, включающий в себя представление 242 базового слоя (одноадресной передачи) и представление 244 улучшающего слоя (широковещательной передачи). Представление 242 базового слоя включает в себя сегменты 246A-246E (сегменты 246), тогда как представление 244 улучшающего слоя включает в себя сегменты 248A, 248B (сегменты 248). Допустим, что короткий сегмент составляет 0,5 секунды по длительности, и частота кадров составляет 50 FPS. На основе технологий, описанных выше, установление и передача сигналов могут заключаться в следующем:

Каждый из сегментов 246, 248 представляет собой сегмент с произвольным доступом (хотя на фиг. 8 показан сегмент 246A, включающий в себя IDR, точка произвольного доступа не обязательно ограничивается IDR, поскольку могут быть предусмотрены другие функциональные точки входа. Открытых GOP может быть достаточно.)

Сегменты 248 в представлении 244 улучшающего слоя (т.е. в представлении широковещательной передачи) в четыре раза превышают по временной длительности сегменты 246 в представлении 242 базового слоя (т.е. в представлении одноадресной передачи).

Передача сигналов может заключаться в следующем для адаптивного набора 240 согласно примеру по фиг. 8:

```

- AdaptationSet
-- @timescale=50
-- Switching: @period=100, @type="media"
--
SegmentTemplate@media="http://example.com/$RepresentationID$/segment_
$Time$.mp4
- Representation: @id=242, @randomAccessPeriod=25
- SegmentTimeline.S: @t=0, @d=25, @r=-1
- Representation: @id=244, @randomAccessPeriod=100,
@dependencyID=242
- SegmentTimeline.S: @t=0, @d=100, @r=-1

```

Фиг. 9 является концептуальной схемой, иллюстрирующей примерный вариант использования, включающий в себя быструю настройку с помощью точки доступа к потоку (SAP) типа 3 в соответствии с технологиями этого раскрытия сущности. В частности, в примере по фиг. 9, адаптивный набор 254 включает в себя представление 250, которое включает в себя сегменты 252A-252E, каждый из которых включает в себя открытую GOP. Хотя не показано на фиг. 9, адаптивный набор 254 может включать в себя представления в дополнение к представлению 250. Допустим, что короткий сегмент составляет 0,5 секунды по длительности, и частота кадров составляет 50 FPS. Передача сигналов может заключаться в следующем для адаптивного набора 254 согласно примеру по фиг. 9:

```

- AdaptationSet
-- @timescale=50
-- @randomAccessPeriod=25
-- SegmentTimeline.S: @t=0, @d=25, @r=-1
--
SegmentTemplate@media="http://example.com/$RepresentationID$/segment_
$Time$.mp4
- Representation: @id=250

```

Фиг. 10 является концептуальной схемой, иллюстрирующей примерный вариант использования, включающий в себя быструю настройку и гибридизацию. В частности, в этом примере, адаптивный набор 260 включает в себя представление 262 и представление 264. Представление 262 включает в себя сегменты 266A-266F (сегменты 266), тогда как представление 264 включает в себя сегменты 268A-268F (сегменты 268). Допустим, что короткий сегмент составляет 0,5 секунды по длительности, и частота кадров составляет 50 FPS. На основе технологий, поясненных выше, установление и передача сигналов могут заключаться в следующем:

Каждый сегмент представляет собой сегмент с произвольным доступом.

Каждый четвертый сегмент представляет собой сегмент с переключением для переключения мультимедиа.

Передача сигналов может заключаться в следующем для адаптивного набора 260 согласно фиг. 10:

```

- AdaptationSet
-- @timescale=50
-- SegmentTimeline.S: @t=0, @d=25, @r=-1
-- @randomAccessPeriod=25
-- Switching: @period=100, @type="media"
--
SegmentTemplate@media="http://example.com/$RepresentationID$/segment_
$Time$.mp4
- Representation: @id=262
- Representation: @id=264

```

Фиг. 11 является концептуальной схемой, иллюстрирующей примерный вариант использования, включающий в себя быструю настройку, гибридизацию и открытые GOP. Предложение сегментов, идентичное предложению сегментов по фиг. 10, показано на фиг. 11. Помимо этого, пример по фиг. 11 иллюстрирует обход 270 сегментов, который представляет сегменты, извлеченные посредством клиентского устройства, такого как клиентское устройство 40 (фиг. 1). Иными словами, клиентское устройство 40 может первоначально извлекать сегмент 266A представления 262, затем переключаться на представление 264 (например, вследствие изменения доступной полосы пропускания сети). Чтобы переключаться, кли-

ентское устройство 40 может извлекать сегмент 268B. В этом примере, сегмент 266A представляет собой IDR-сегмент, тогда как сегмент 268B представляет собой сегмент открытой GOP. В соответствии с технологиями этого раскрытия сущности, поскольку сегмент 268B представляет собой сегмент открытой GOP, клиентское устройство 40 может осуществлять переключение на 268B без ожидания IDR-сегмента (например, сегмента 268E) представления 264. Клиентское устройство 40 также извлекает сегмент 268C представления 264. Затем, клиентское устройство 40 снова переключает представления, на этот раз на представление 262, извлекая сегмент 266D, который также представляет собой сегмент открытой GOP. В этом примере, клиентское устройство 40 извлекает сегменты 266E и 266F из представления 262, согласно обходу 270 сегментов.

Переключение может возникать в SAP типа 3. Допустим, что короткий сегмент составляет 0,5 секунды по длительности, и частота кадров составляет 50 FPS. На основе технологий, поясненных выше, установление и передача сигналов могут заключаться в следующем:

Каждый сегмент представляет собой сегмент с произвольным доступом.

Каждый четвертый сегмент представляет собой сегмент с переключением для переключения мультимедиа.

Каждый сегмент представляет собой сегмент с переключением для переключения открытых GOP.

Передача сигналов может заключаться в следующем для адаптивного набора 260 согласно фиг. 11:

```
- AdaptationSet
-- @timescale=50
-- SegmentTimeline.S: @t=0, @d=25, @r=-1
-- @randomAccessPeriod=25
-- Switching: @period=100, @type="media"
-- Switching: @period=25, @type="open GOP"
--

SegmentTemplate@media="http://example.com/$RepresentationID
$/segment_ $Time$.mp4
- Representation: @id=262
- Representation: @id=264
```

Фиг. 12 является концептуальной схемой, иллюстрирующей другой примерный вариант использования, включающий в себя быструю настройку и гибридизацию с открытыми GOP. В этом примере, адаптивный набор 280 включает в себя представление 282 одноадресной передачи и представление 284 ширококвещательной передачи. Представление 282 одноадресной передачи включает в себя сегменты 286A-286F (сегменты 286), тогда как представление 284 ширококвещательной передачи включает в себя сегменты 288A, 288B (сегменты 288). Допустим, что короткий сегмент составляет 0,5 секунды по длительности, и частота кадров составляет 50 FPS. На основе технологий, поясненных выше, установление и передача сигналов могут заключаться в следующем:

Каждый сегмент представляет собой сегмент с произвольным доступом.

Сегменты 288 в представлении 284 ширококвещательной передачи в 4 раза превышают по временной длительности сегменты 286 в представлении 282 одноадресной передачи.

Сегмент в позициях перекрытия ширококвещательной передачи/одноадресной передачи (например, сегменты 286A, 286E, 288A, 288B) представляет собой сегменты с переключением.

Передача сигналов может заключаться в следующем для адаптивного набора 280 согласно фиг. 12:

```
- AdaptationSet
-- @timescale=50
-- Switching: @period=100, @type="media"
--

SegmentTemplate@media="http://example.com/$RepresentationID
$/segment_ $Time$.mp4
- Representation: @id=282, @randomAccessPeriod=100
- SegmentTimeline.S: @t=0, @d=100, @r=-1
- Representation: @id=284, @randomAccessPeriod=25
- SegmentTimeline.S: @t=0, @d=25, @r=-1
```

Фиг. 13 является концептуальной схемой, иллюстрирующей примерный вариант использования, включающий в себя быструю настройку и очень низкую задержку. В этом примере, адаптивный набор 290 включает в себя представление 292 одноадресной передачи и представление 294 ширококвещательной передачи. Представление 292 одноадресной передачи включает в себя сегменты 296A-296F (сегменты 296), тогда как представление 294 ширококвещательной передачи включает в себя сегменты 298A, 298B

(сегменты 298). Допустим, что короткий сегмент составляет 0,5 секунды по длительности, и частота кадров составляет 50 FPS. На основе технологий, поясненных выше, установление и передача сигналов могут заключаться в следующем:

Каждый сегмент представляет собой сегмент с произвольным доступом.

Сегменты 298 в представлении 294 ширококвещательной передачи в 4 раза превышают по временной длительности сегменты 296 в представлении 292 одноадресной передачи.

Сегмент в позициях перекрытия ширококвещательной передачи/одноадресной передачи (например, сегменты 296А, 296Е, 298А, 298В) представляет собой сегменты с переключением.

Кроме того, не все сегменты 296 представления 292 обеспечивают информацию для переключения. Например, сегмент 296С обеспечивает возможность переключения с представления 294 ширококвещательной передачи на представление 292 одноадресной передачи (например, если ширококвещательная услуга становится недоступной). Тем не менее, сегменты 296В, 296D и 296F соответствуют формату мультимедийных сегментов единиц доставки и не включают в себя точки переключения. Это обеспечивает возможность выделения большего числа битов сегментам 296В, 296D и 296F не внутренне прогнозированным кадрам (например, взаимно прогнозированным кадрам), например, таким образом, что эти кадры могут кодироваться с более высоким качеством.

Передача сигналов может заключаться в следующем для адаптивного набора 290 согласно фиг. 13:

```
- AdaptationSet
-- @timescale=50
-- Switching: @period=100, @type="media"
--
SegmentTemplate@media="http://example.com/$RepresentationID
$/segment_ $Time$.mp4
- Representation: @id=292, @randomAccessPeriod=100
- SegmentTimeline.S: @t=0, @d=100, @r=-1
- Representation: @id=294, @randomAccessPeriod=50
- SegmentTimeline.S: @t=0, @d=25, @r=-1
```

Фиг. 14 является концептуальной схемой, иллюстрирующей другой примерный вариант использования, включающий в себя быструю настройку и очень низкую задержку. В этом примере, адаптивный набор 300 включает в себя представление 302 и представление 304. Представление 302 включает в себя сегменты 306А-306F (сегменты 306), тогда как представление 304 включает в себя сегменты 308А-308F (сегменты 308). Допустим, что короткий сегмент составляет 0,5 секунды по длительности, и частота кадров составляет 50 FPS. На основе технологий, поясненных выше, установление и передача сигналов могут заключаться в следующем:

Каждый из сегментов 306 в представлении 302 представляет собой сегмент с произвольным доступом.

Иными словами, как показано на фиг. 14, каждый из сегментов 306 включает в себя IDR-изображение. Тем не менее, сегменты 308А и 308Е представления 304 включают в себя IDR-изображения, тогда как сегменты 308В, 308С, 308D и 308F не включают в себя IDR-изображения. Это обеспечивает возможность клиентскому устройству, к примеру, клиентскому устройству 40 (Фиг. 1), быстро настраиваться на мультимедийный контент адаптивного набора 300 посредством извлечения последнего доступного мультимедийного контента из сегментов 306, а затем переключения на представление 304, когда следующий из сегментов 308, включающих в себя IDR, доступен.

Передача сигналов может заключаться в следующем для адаптивного набора 300 согласно фиг. 14:

```
- AdaptationSet
-- @timescale=50
-- Switching: @period=100, @type="media"
-- SegmentTimeline.S: @t=0, @d=25, @r=-1
--
SegmentTemplate@media="http://example.com/$RepresentationID
$/segment_ $Time$.mp4
- Representation: @id=302, @randomAccessPeriod=25
- Switching: @period=25, @type="media"
- Representation: @id=304, @randomAccessPeriod=100
- Switching: @period=100, @type="media"
```

Таким образом, технологии этого раскрытия сущности включают в себя:

Дополнительные новые типы сегментов.

Дополнительную передачу MPD-сигналов для переключения и @randomAccessPeriod.

Определения для различных типов переключения.

Переключение мультимедиа: Совмещение сегментов и SAP-тип 1 или 2.

Переключение потоков битов: конкатенация разрешается. Переключение открытых GOP.

Добавление профиля, который документирует расширения и ограничения.

Документирование всех проблем касательно обратной совместимости.

Предоставление более подробных примеров.

Остаются нерешенные вопросы и альтернативы. Следующие проблемы остаются открытыми:

Возможна передача сигналов на основе числа в качестве добавления или альтернативы технологиям этого раскрытия сущности, которая может обеспечивать определенные импликации и преимущества.

Также возможны различные типы переключения открытых GOP в качестве добавления или альтернативы технологиям этого раскрытия сущности, которые могут параллелизовать повторную дискретизацию и неповторную дискретизацию.

Дополнительные или альтернативные форматы мультимедиа могут использоваться относительно форматов мультимедиа, поясненных выше.

Субсегменты, в дополнение или альтернативно полным сегментам, также могут использоваться в некоторых примерах. Поле индексов сегментов (SIDX), к примеру, показанное на фиг. 6 выше, может передавать в сигналах местоположения субсегментов, и/или дополнительная информация может передаваться в сигналах (например, в метаданных файла и/или в файле манифеста, таком как в MPD).

Фиг. 15 является блок-схемой последовательности операций, иллюстрирующей примерный способ для извлечения сегмента представления мультимедийного контента в соответствии с технологиями этого раскрытия сущности. Способ по фиг. 15 описывается как осуществляемый посредством серверного устройства 60 и клиентского устройства 40 по фиг. 4. Тем не менее, следует понимать, что способ может осуществляться посредством других устройств. Например, весь или части способа, приписанные серверному устройству, могут осуществляться посредством устройства 20 подготовки контента по фиг. 4 (например, в дополнение или альтернативно серверному устройству 60 по фиг. 4). Аналогично, весь или части способа, приписанные клиентскому устройству, могут осуществляться посредством промежуточного программного модуля клиентского устройства, которое выполнено с возможностью принимать мультимедийные данные через широковещательную передачу и/или одноадресную передачу.

В этом примере, серверное устройство 60 первоначально принимает кодированный мультимедийный поток (320). В некоторых примерах, серверное устройство 60 принимает кодированный мультимедийный поток из устройства 20 подготовки контента, тогда как в других примерах, серверное устройство 60 может включать в себя один или более кодеров для того, чтобы кодировать необработанные мультимедийные данные, чтобы формировать кодированный мультимедийный поток.

Серверное устройство 60 затем, в этом примере, определяет типы и местоположения сегментов в кодированном мультимедийном потоке (322). В некоторых примерах, серверное устройство 60 может формировать сегменты (т.е. независимо извлекаемые файлы), тогда как в других примерах, серверное устройство 60 может принимать и анализировать сегменты в качестве части кодированного мультимедийного потока и определять типы для сегментов на основе их характеристик. Выше пояснены характеристики различных типов сегментов, таких как мультимедийные сегменты единиц доставки, мультимедийные сегменты с произвольным доступом, сегменты без перекрытия и мультимедийные сегменты с переключением. Таким образом, серверное устройство 60 может анализировать каждый сегмент, чтобы определять то, какие из этих типов сегментов совпадают с характеристиками анализируемого сегмента. Кроме того, серверное устройство 60 может определять местоположения сегментов каждого типа в кодированном мультимедийном потоке. Например, серверное устройство 60 может определять частоты, с которыми возникает каждый тип сегмента. В качестве примера, относительно фиг. 7, сегменты, включающие в себя IDR (т.е. мультимедийные сегменты с произвольным доступом), возникают каждый четвертый сегмент каждого из представлений 232, 234.

В этом примере, серверное устройство 60 затем составляет файл манифеста (к примеру, MPD), передающий в сигналах типы и местоположения сегментов (324). Альтернативно, серверное устройство 60 может принимать файл манифеста, частично или полностью составленный согласно технологиям этого раскрытия сущности, из устройства 20 подготовки контента. Серверное устройство 60 может составлять файл манифеста таким образом, что он включает в себя информацию (т.е. "передает в сигналах") типов и местоположений сегментов в соответствующем представлении каждого адаптивного набора, представленного посредством файла манифеста. Серверное устройство 60 может составлять файл манифеста таким образом, что он включает в себя данные, аналогичные данным, поясненным выше относительно примеров фиг. 7-14. Следует понимать, что файл манифеста является отдельным от представлений и мультимедийных данных самих представлений. Например, файл манифеста может быть доступным для запроса отдельно от запросов, выполняемых на предмет мультимедийных данных (например, сегментов или частей сегментов), описанных посредством файла манифеста.

Серверное устройство 60 затем может выводить файл манифеста (326), например, в клиентское устройство 40. В некоторых примерах, клиентское устройство 40 может первоначально запрашивать файл манифеста, например, через одноадресный запрос для файла манифеста. В других примерах, клиентское

устройство 40 может подписываться на широковещательную передачу, и серверное устройство 60 может периодически выводить файл манифеста через широковещательную передачу. В любом случае, клиентское устройство 40 может принимать файл манифеста (328), который выведен посредством серверного устройства 60.

Клиентское устройство 40 затем может определять типы и местоположения сегментов из файла манифеста (330). Например, клиентское устройство 40 может определять то, что файл манифеста указывает то, что конкретный адаптивный набор включает в себя представления, включающие в себя, например, мультимедийные сегменты единиц доставки, мультимедийные сегменты с произвольным доступом, сегменты без перекрытия и мультимедийные сегменты с переключением. Клиентское устройство 40 также может определять местоположения каждого из этих типов сегментов. Например, клиентское устройство 40 может определять частоты, с которыми все или часть этих типов сегментов возникают из файла манифеста.

Клиентское устройство 40 может определять одно из представлений, с которого можно начинать извлечение мультимедийных данных. Клиентское устройство 40 может выполнять любой из различных вариантов использования, поясненных выше. Чтобы достигать воспроизведения с низкой задержкой, клиентское устройство 40 может определять то, какое из представлений, если таковые имеются, имеет самые частые сегменты, включающие в себя точки доступа к потоку (SAP), например, IDR-кадры. Такое представление может включать в себя сегменты, доступные для извлечения через одноадресную передачу. Клиентское устройство 40 может быть выполнено с возможностью первоначально извлекать такие сегменты из представления одноадресной передачи, затем переключаться на представление широковещательной передачи в следующей доступной SAP представления широковещательной передачи (снова, как указано посредством файла манифеста).

В любом случае, клиентское устройство 40 может определять сегмент представления, обеспечивающего начальную точку (332). Как пояснено выше, сегмент может содержать мультимедийный сегмент с произвольным доступом, т.е. соответствовать формату мультимедийных сегментов с произвольным доступом. Аналогично, клиентское устройство 40 может определять универсальный указатель ресурса (URL-адрес) для определенного сегмента, например, согласно шаблону, указываемому посредством файла манифеста. Клиентское устройство 40 затем может запрашивать определенный сегмент (334), например, посредством выдачи HTTP GET- или частичного GET-запроса на предмет URL-адреса для серверного устройства 60.

Серверное устройство 60 затем может принимать запрос (336) и после этого отправлять запрашиваемый сегмент в клиентское устройство 40 (338) в ответ на запрос. После приема сегмента (340), клиентское устройство 40 может первоначально буферизовать данные принимаемого сегмента, затем в конечном счете декодировать и представлять данные принимаемого сегмента (342).

Как пояснено выше, после начального извлечения определенного сегмента представления, клиентское устройство 40, может определять то, следует или нет и когда следует переключаться на другое представление. Например, начальное представление может включать в себя очень частые SAP, и целевое представление может включать в себя относительно нечастые SAP. Клиентское устройство 40 может продолжать запрашивать сегменты из начального представления до достижения сегмента, включающего в себя SAP (например, мультимедийного сегмента с произвольным доступом или мультимедийного сегмента с переключением) целевого представления. Затем клиентское устройство 40 может либо начинать запрос сегментов с целевого представления (если целевое представление доступно через одноадресную передачу), либо подписываться на широковещательную услугу, которая транспортирует мультимедийные данные целевого представления (если целевое представление доступно через широковещательную передачу).

Таким образом, фиг. 15 представляет пример способа, включающего в себя определение, из файла манифеста, множества типов сегментов, включенных в представление мультимедийного контента, причем одна или более функций обеспечиваются посредством каждого из типов сегментов и позиций сегментов, соответствующих каждому из типов сегментов в представлении, при этом, по меньшей мере, один из типов сегментов обеспечивает точку, в которой можно начинать извлечение данных из представления, определение, из файла манифеста, сегмента представления, соответствующего типу, который обеспечивает точку, в которой можно начинать извлечение данных из представления, и извлечение определенного сегмента из представления.

Фиг. 15 также представляет пример способа, включающего в себя составление файла манифеста, указывающего множество типов сегментов, включенных в представление мультимедийного контента, причем одна или более функций обеспечиваются посредством каждого из типов сегментов, причем позиции сегментов соответствуют каждому из типов сегментов в представлении, при этом, по меньшей мере, один из типов сегментов обеспечивает точку, в которой можно начинать извлечение данных из представления, и сегмент представления, соответствующего типу, который обеспечивает точку, в которой можно начинать извлечение данных из представления, отправку файла манифеста в клиентское устройство, и в ответ на запрос из клиентского устройства на предмет сегмента, соответствующего типу, который обеспечивает точку, в которой можно начинать извлечение данных из представления, отправку

сегмента, который обеспечивает точку, в которой можно начинать извлечение данных из представления, в клиентское устройство.

В одном или более примеров, описанные функции могут быть реализованы в аппаратных средствах, программном обеспечении, микропрограммном обеспечении или любой комбинации вышеозначенного. При реализации в программном обеспечении, функции могут быть сохранены или переданы, в качестве одной или более инструкций или кода, по машиночитаемому носителю и выполнены посредством аппаратного модуля обработки. Машиночитаемые носители могут включать в себя машиночитаемые носители хранения данных, которые соответствуют материальному носителю, такие как носители хранения данных, или среды связи, включающие в себя любой носитель, который упрощает перенос компьютерной программы из одного места в другое, например, согласно протоколу связи. Таким образом, машиночитаемые носители, в общем, могут соответствовать (1) материальному машиночитаемому носителю хранения данных, который является энергонезависимым, или (2) среде связи, такой как сигнал или несущая. Носители хранения данных могут представлять собой любые доступные носители, к которым может осуществляться доступ посредством одного или более компьютеров или одного или более процессоров, с тем чтобы извлекать инструкции, код и/или структуры данных для реализации технологий, описанных в этом раскрытии сущности. Компьютерный программный продукт может включать в себя машиночитаемый носитель.

В качестве примера, а не ограничения, эти машиночитаемые носители хранения данных могут содержать RAM, ROM, EEPROM, CD-ROM или другое устройство хранения данных на оптических дисках, устройство хранения данных на магнитных дисках или другие магнитные устройства хранения, флэш-память либо любой другой носитель, который может быть использован для того, чтобы сохранять требуемый программный код в форме инструкций или структур данных, и к которому можно осуществлять доступ посредством компьютера. Кроме того, любое соединение корректно называть машиночитаемым носителем. Например, если инструкции передаются из веб-узла, сервера или другого удаленного источника с помощью коаксиального кабеля, оптоволоконного кабеля, "витой пары", цифровой абонентской линии (DSL) или беспроводных технологий, таких как инфракрасные, радиопередающие и микроволновые среды, то коаксиальный кабель, оптоволоконный кабель, "витая пара", DSL или беспроводные технологии, такие как инфракрасные, радиопередающие и микроволновые среды, включаются в определение носителя. Тем не менее, следует понимать, что машиночитаемые носители хранения данных и носители хранения данных не включают в себя соединения, несущие, сигналы или другие энергозависимые носители, а вместо этого направлены на энергонезависимые материальные носители хранения данных. Диск (disk) и диск (disc) при использовании в данном документе включают в себя компакт-диск (CD), лазерный диск, оптический диск, универсальный цифровой диск (DVD), гибкий диск и Blu-Ray-диск, при этом диски (disk) обычно воспроизводят данные магнитно, тогда как диски (disc) обычно воспроизводят данные оптически с помощью лазеров. Комбинации вышеперечисленного также следует включать в число машиночитаемых носителей.

Инструкции могут выполняться посредством одного или более процессоров, например, одного или более процессоров цифровых сигналов (DSP), микропроцессоров общего назначения, специализированных интегральных схем (ASIC), программируемых пользователем вентильных матриц (FPGA) либо других эквивалентных интегральных или дискретных логических схем. Соответственно, термин "процессор" при использовании в данном документе может означать любую вышеуказанную структуру или другую структуру, подходящую для реализации технологий, описанных в данном документе. Помимо этого, в некоторых аспектах функциональность, описанная в данном документе, может быть обеспечена в рамках специализированных аппаратных и/или программных модулей, выполненных с возможностью кодирования или декодирования либо встроенных в комбинированный кодек. Кроме того, технологии могут быть полностью реализованы в одной или более схем или логических элементов.

Технологии этого раскрытия сущности могут быть реализованы в широком спектре устройств или приборов, в том числе в беспроводном переносном телефоне, в интегральной схеме (IC) или в наборе IC (к примеру, в наборе микросхем). Различные компоненты, модули или блоки описываются в этом раскрытии сущности для того, чтобы подчеркивать функциональные аспекты устройств, выполненных с возможностью осуществлять раскрытые технологии, но необязательно требуют реализации посредством различных аппаратных модулей. Наоборот, как описано выше, различные блоки могут быть комбинированы в аппаратный модуль кодека или обеспечены посредством набора взаимодействующих аппаратных модулей, включающих в себя один или более процессоров, как описано выше, в сочетании с надлежащим программным обеспечением и/или микропрограммным обеспечением.

Описаны различные примеры. Эти и другие примеры находятся в пределах объема прилагаемой формулы изобретения.

ФОРМУЛА ИЗОБРЕТЕНИЯ

1. Способ извлечения мультимедийных данных из серверного устройства, при этом упомянутый способ содержит этапы, на которых:

принимают от серверного устройства файл манифеста, содержащий информацию, указывающую какой формат из множества форматов мультимедийных сегментов соответствует мультимедийным сегментам, включенным в представление мультимедийного контента, при этом упомянутое представление мультимедийного контента содержится в адаптивном наборе, являющемся набором представлений мультимедийного контента, адаптированных в отношении по меньшей мере скорости передачи битов, частоты кадров, пространственного разрешения, каждое из представлений мультимедийного контента представляет собой структурированную совокупность данных, которая является доступной для клиентских устройств потоковой передачи и содержит мультимедийные данные кинофрагментов, множество форматов мультимедийных сегментов включает в себя:

формат мультимедийных сегментов единиц доставки, который указывает то, что соответствующий мультимедийный сегмент является единицей доставки, включающей в себя, по меньшей мере, один кинофрагмент и предназначенной для доставки в клиентское устройство;

формат мультимедийных сегментов с произвольным доступом, который указывает то, что соответствующие мультимедийные сегменты содержат точки доступа к потоку (SAP), которые являются указателями в представлении мультимедийного контента, обеспечивающими возможность начать воспроизведение медиапотока, в которых начинается извлечение мультимедийных данных из представления мультимедийного контента;

формат мультимедийных сегментов без перекрытия, который указывает то, что соответствующие мультимедийные сегменты не перекрывают времена начала и завершения других мультимедийных сегментов в представлении мультимедийного контента и других мультимедийных сегментов в других представлениях мультимедийного контента в адаптивном наборе; и

формат мультимедийных сегментов с переключением, который указывает то, что соответствующие мультимедийные сегменты содержат точки доступа к потоку (SAP), которые являются указателями в представлении мультимедийного контента, в которых извлечение мультимедийных данных из других представлений мультимедийного контента в адаптивном наборе переключается без повторной инициализации для доступа к упомянутым соответствующим мультимедийным сегментам на извлечение мультимедийных данных из упомянутого представления мультимедийного контента с соответствующих мультимедийных сегментов;

определяют из упомянутой информации, содержащейся в файле манифеста, какой формат из множества форматов мультимедийных сегментов соответствует мультимедийным сегментам, включенным в представление мультимедийного контента;

используют упомянутые определенные форматы для определения мультимедийного сегмента, содержащего SAP, запрашивают у серверного устройства мультимедийный сегмент, содержащий SAP, и начинают извлечение мультимедийных данных из представления мультимедийного контента с мультимедийного сегмента, содержащего SAP, из серверного устройства; и

используют упомянутые определенные форматы для определения мультимедийного сегмента, имеющего формат мультимедийных сегментов с переключением, запрашивают у серверного устройства мультимедийный сегмент, имеющий формат мультимедийных сегментов с переключением, и переключают извлечение мультимедийных данных с другого представления мультимедийного контента в адаптивном наборе на извлечение мультимедийных данных из представления мультимедийного контента, содержащего определенный мультимедийный сегмент, имеющий формат мультимедийных сегментов с переключением, с определенного мультимедийного сегмента, имеющего формат мультимедийных сегментов с переключением.

2. Способ передачи в сигналах мультимедийной информации для извлечения мультимедийных сегментов мультимедийного контента из серверного устройства, при этом упомянутый способ содержит этапы, на которых:

составляют файл манифеста, указывающий какой формат из множества форматов мультимедийных сегментов соответствует мультимедийным сегментам, включенным в представление мультимедийного контента, при этом упомянутое представление мультимедийного контента содержится в адаптивном наборе, являющемся набором представлений мультимедийного контента, адаптированных в отношении по меньшей мере скорости передачи битов, частоты кадров, пространственного разрешения, каждое из представлений мультимедийного контента представляет собой структурированную совокупность данных, которая является доступной для клиентских устройств потоковой передачи и содержит мультимедийные данные кинофрагментов, множество форматов мультимедийных сегментов включает в себя:

формат мультимедийных сегментов единиц доставки, который указывает то, что соответствующий мультимедийный сегмент является единицей доставки, включающей в себя, по меньшей мере, один кинофрагмент и предназначенной для доставки в клиентское устройство;

формат мультимедийных сегментов с произвольным доступом, который указывает то, что соответствующие сегменты содержат точки доступа к потоку (SAP), которые являются указателями в представлении мультимедийного контента, обеспечивающими возможность начать воспроизведение медиапотока, в которых начинается извлечение мультимедийных данных из представления мультимедийного контента;

формат мультимедийных сегментов без перекрытия, который указывает то, что соответствующие мультимедийные сегменты не перекрывают времена начала и завершения других мультимедийных сегментов в представлении мультимедийного контента и других мультимедийных сегментов в других представлениях мультимедийного контента в адаптивном наборе; и

формат мультимедийных сегментов с переключением, который указывает то, что соответствующие мультимедийные сегменты содержат точки доступа к потоку (SAP), которые являются указателями в представлении мультимедийного контента, в которых извлечение мультимедийных данных из других представлений мультимедийного контента в адаптивном наборе переключается без повторной инициализации для доступа к упомянутым соответствующим мультимедийным сегментам на извлечение мультимедийных данных из упомянутого представления мультимедийного контента с упомянутых соответствующих мультимедийных сегментов;

отправляют файл манифеста в клиентское устройство; и

в ответ на запрос из клиентского устройства мультимедийного сегмента, соответствующего одному из упомянутого множества форматов мультимедийного сегмента, отправляют мультимедийный сегмент, который соответствует упомянутому формату мультимедийного сегмента, в клиентское устройство, при этом мультимедийный сегмент, содержащий SAP, отправляют для начала извлечения мультимедийных данных из представления мультимедийного контента, и мультимедийный сегмент, имеющий формат мультимедийных сегментов с переключением, отправляют для переключения извлечения мультимедийных данных с другого представления мультимедийного контента в адаптивном наборе на извлечение мультимедийных данных из представления мультимедийного контента, содержащего мультимедийный сегмент, имеющий формат мультимедийных сегментов с переключением.

3. Способ по п.1 или 2, в котором мультимедийный сегмент, соответствующий формату мультимедийных сегментов единиц доставки, содержит значение `dums` в поле типов сегментов для мультимедийного сегмента;

каждый из автономных кинофрагментов содержит поле кинофрагментов (`moof`) и поле мультимедийных данных (`mdat`), которое содержит мультимедийные выборки, которые не используют внешние ссылки на данные, к которым обращаются посредством дорожки в поле кинофрагментов; каждое из полей `moof` содержит, по меньшей мере, один фрагмент дорожки; каждое из полей `moof` не использует внешние ссылки; флаг `default-base-is-moof` мультимедийного сегмента задается как истинный; и флаг `base-data-offset-present` мультимедийного сегмента задается как ложный.

4. Способ по п.1 или 2, в котором мультимедийный сегмент, соответствующий формату мультимедийных сегментов с произвольным доступом, содержит сегменты с произвольным доступом, ординальная первая единица доступа в каждом кинофрагменте мультимедийных сегментов соответствует I_{SAU} точки доступа к потоку (SAP) типа 1, 2 или 3; и включает в себя всю необходимую информацию для того, чтобы осуществлять доступ к мультимедийным данным в потоке битов после мультимедийных сегментов.

5. Способ по п.4, в котором мультимедийный сегмент, соответствующий формату мультимедийных сегментов с произвольным доступом, содержит, по меньшей мере, одно из изображения на основе мгновенного обновления декодера (IDR), изображения на основе доступа к нерабочей ссылке (BLA) или изображения на основе чистого произвольного доступа (CRA).

6. Способ по п.1 или 2, в котором мультимедийный сегмент, соответствующий формату мультимедийных сегментов с произвольным доступом, включает в себя одно или более полей индексов сегментов (`sidx`), и ординальное первое поле `sidx` предшествует всем полям `moof` мультимедийного сегмента и описывает весь мультимедийный сегмент.

7. Способ по п.1 или 2, в котором формат мультимедийных сегментов с переключением указывает то, что ординальная первая выборка в ординальном первом кинофрагменте соответствующих мультимедийных сегментов соответствует I_{SAU} точки доступа к потоку (SAP) типа 1 или 2.

8. Клиентское устройство для извлечения мультимедийных данных из серверного устройства, причем упомянутое клиентское устройство содержит:

средство для приема от серверного устройства файла манифеста, содержащего информацию, указывающую какой формат из множества форматов мультимедийных сегментов соответствует мультимедийным сегментам, включенным в представление мультимедийного контента, при этом упомянутое представление мультимедийного контента содержится в адаптивном наборе, являющемся набором представлений мультимедийного контента, адаптированных в отношении по меньшей мере скорости передачи битов, частоты кадров, пространственного разрешения, каждое из представлений мультимедийного контента представляет собой структурированную совокупность данных, которая является доступной для клиентских устройств потоковой передачи и содержит мультимедийные данные кинофрагментов, мно-

жество форматов мультимедийных сегментов включает в себя:

формат мультимедийных сегментов единиц доставки, который указывает то, что соответствующий мультимедийный сегмент является единицей доставки, включающей в себя, по меньшей мере, один кинофрагмент и предназначенной для доставки в клиентское устройство;

формат мультимедийных сегментов с произвольным доступом, который указывает то, что соответствующие сегменты содержат точки доступа к потоку (SAP), которые являются указателями в представлении мультимедийного контента, обеспечивающими возможность начать воспроизведение медиапотока, в которых начинается извлечение мультимедийных данных из представления мультимедийного контента;

формат мультимедийных сегментов без перекрытия, который указывает то, что соответствующие мультимедийные сегменты не перекрывают времена начала и завершения других мультимедийных сегментов в представлении мультимедийного контента и других мультимедийных сегментов в других представлениях мультимедийного контента в адаптивном наборе; и

формат мультимедийных сегментов с переключением, который указывает то, что соответствующие мультимедийные сегменты содержат точки доступа к потоку (SAP), которые являются указателями в представлении мультимедийного контента, в которых извлечение мультимедийных данных из других представлений мультимедийного контента в адаптивном наборе переключается без повторной инициализации для доступа к упомянутым соответствующим мультимедийным сегментам на извлечение мультимедийных данных из представления мультимедийного контента с упомянутых соответствующих мультимедийных сегментов;

средство для определения из упомянутой информации, содержащейся в файле манифеста, какой формат из множества форматов мультимедийных сегментов соответствует мультимедийным сегментам, включенным в представление мультимедийного контента; и

средство для использования упомянутых определенных форматов для определения мультимедийного сегмента, содержащего SAP, запроса у серверного устройства мультимедийного сегмента, содержащего SAP, и начала извлечения мультимедийных данных из представления мультимедийного контента с мультимедийного сегмента, содержащего SAP, из серверного устройства; и для определения мультимедийного сегмента, имеющего формат мультимедийных сегментов с переключением, запроса у серверного устройства мультимедийного сегмента, имеющего формат мультимедийных сегментов с переключением, и переключения извлечения мультимедийных данных с другого представления мультимедийного контента в адаптивном наборе на извлечение мультимедийных данных из представления мультимедийного контента, содержащего определенный мультимедийный сегмент, имеющий формат мультимедийных сегментов с переключением, с определенного мультимедийного сегмента, имеющего формат мультимедийных сегментов с переключением.

9. Серверное устройство для передачи в сигналах мультимедийной информации для извлечения мультимедийных сегментов мультимедийного контента из серверного устройства, причем упомянутое серверное устройство содержит:

средство для составления файла манифеста, указывающего какой формат из множества форматов мультимедийных сегментов соответствует мультимедийным сегментам, включенным в представление мультимедийного контента, при этом упомянутое представление мультимедийного контента содержится в адаптивном наборе, являющемся набором представлений мультимедийного контента, адаптированных в отношении по меньшей мере скорости передачи битов, частоты кадров, пространственного разрешения, каждое из представлений мультимедийного контента представляет собой структурированную совокупность данных, которая является доступной для клиентских устройств потоковой передачи и содержит мультимедийные данные кинофрагментов, множество форматов мультимедийных сегментов включает в себя:

формат мультимедийных сегментов единиц доставки, который указывает то, что соответствующий мультимедийный сегмент является единицей доставки, включающей в себя, по меньшей мере, один кинофрагмент и предназначенной для доставки в клиентское устройство;

формат мультимедийных сегментов с произвольным доступом, который указывает то, что соответствующие мультимедийные сегменты содержат точки доступа к потоку (SAP), которые являются указателями в представлении мультимедийного контента, обеспечивающими возможность начать воспроизведение медиапотока, в которых начинается извлечение мультимедийных данных из представления мультимедийного контента;

формат мультимедийных сегментов без перекрытия, который указывает то, что соответствующие мультимедийные сегменты не перекрывают времена начала и завершения других мультимедийных сегментов в представлении и других мультимедийных сегментов в других представлениях мультимедийного контента в адаптивном наборе; и

формат мультимедийных сегментов с переключением, который указывает то, что соответствующие мультимедийные сегменты содержат точки доступа к потоку (SAP), которые являются указателями в представлении мультимедийного контента, в которых извлечение мультимедийных данных из других представлений мультимедийного контента в адаптивном наборе переключается без повторной инициализации.

зации для доступа к упомянутым соответствующим мультимедийным сегментам на извлечение мультимедийных данных из представления мультимедийного контента с упомянутых соответствующих мультимедийных сегментов;

средство для отправки файла манифеста в клиентское устройство; и

средство для отправки в ответ на запрос из клиентского устройства мультимедийного сегмента, соответствующего одному из упомянутого множества форматов мультимедийного сегмента, мультимедийного сегмента, который соответствует упомянутому формату мультимедийного сегмента, в клиентское устройство, при этом мультимедийный сегмент, содержащий SAP, отправляется для начала извлечения мультимедийных данных из представления мультимедийного контента, и мультимедийный сегмент, имеющий формат мультимедийных сегментов с переключением, отправляется для переключения извлечения мультимедийных данных с другого представления мультимедийного контента в адаптивном наборе на извлечение мультимедийных данных из представления мультимедийного контента, содержащего мультимедийный сегмент, имеющий формат мультимедийных сегментов с переключением.

10. Устройство по п.8 или 9, в котором мультимедийный сегмент, соответствующий формату мультимедийных сегментов единиц доставки, содержит значение *dums* в поле типов сегментов для мультимедийного сегмента;

каждый из автономных кинофрагментов содержит поле кинофрагментов (*moof*) и поле мультимедийных данных (*mdat*), которое содержит мультимедийные выборки, которые не используют внешние ссылки на данные, к которым обращаются посредством дорожки в поле кинофрагментов; каждое из полей *moof* содержит, по меньшей мере, один фрагмент дорожки; каждое из полей *moof* не использует внешние ссылки; флаг *default-base-is-moof* мультимедийного сегмента задается как истинный; и флаг *base-data-offset-present* мультимедийного сегмента задается как ложный.

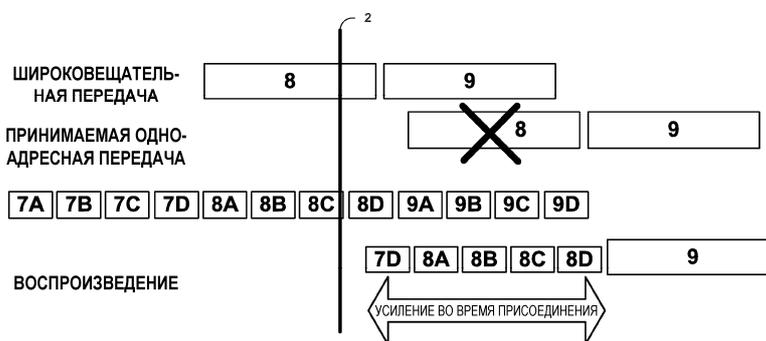
11. Устройство по п.8 или 9, в котором мультимедийный сегмент, соответствующий формату мультимедийных сегментов с произвольным доступом, содержит сегменты с произвольным доступом, ординальная первая единица доступа в каждом кинофрагменте мультимедийных сегментов соответствует I_{SAU} точки доступа к потоку (SAP) типа 1, 2 или 3; и включает в себя всю необходимую информацию для того, чтобы осуществлять доступ к мультимедийным данным в потоке битов после мультимедийных сегментов.

12. Устройство по п.8 или 9, в котором мультимедийный сегмент, соответствующий формату мультимедийных сегментов с произвольным доступом, содержит, по меньшей мере, одно из изображения на основе мгновенного обновления декодера (IDR), изображения на основе доступа к нерабочей ссылке (BLA) или изображения на основе чистого произвольного доступа (CRA).

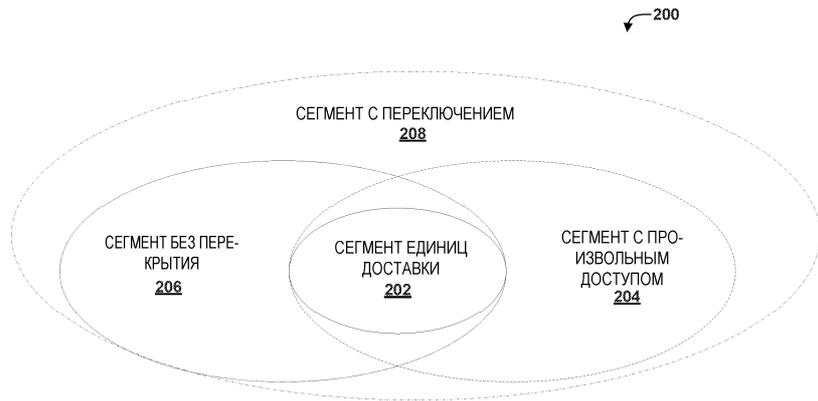
13. Устройство по п.8 или 9, в котором мультимедийный сегмент, соответствующий формату мультимедийных сегментов с произвольным доступом, включает в себя одно или более полей индексов сегментов (*sidx*), и ординальное первое поле *sidx* предшествует всем полям *moof* мультимедийного сегмента и описывает весь мультимедийный сегмент.

14. Устройство по п.8 или 9, в котором формат мультимедийных сегментов с переключением указывает то, что ординальная первая выборка в ординальном первом кинофрагменте соответствующих мультимедийных сегментов соответствует I_{SAU} точки доступа к потоку (SAP) типа 1 или 2.

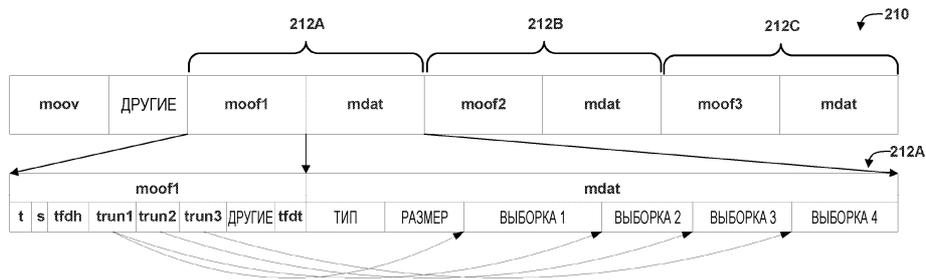
15. Машиночитаемый носитель хранения данных, имеющий сохраненные на нем инструкции, которые при выполнении инструктируют процессору выполнять способ по любому из пп.1-7.



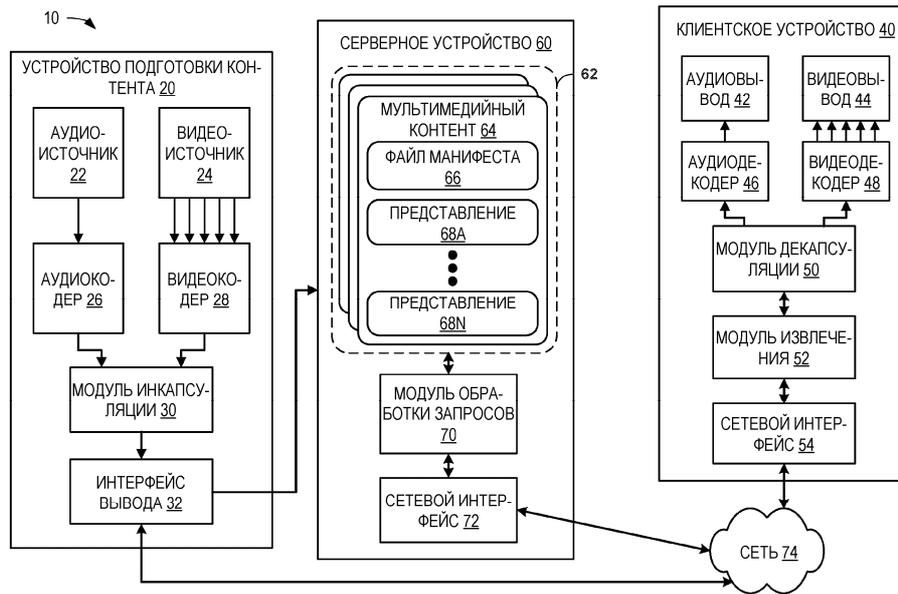
Фиг. 1



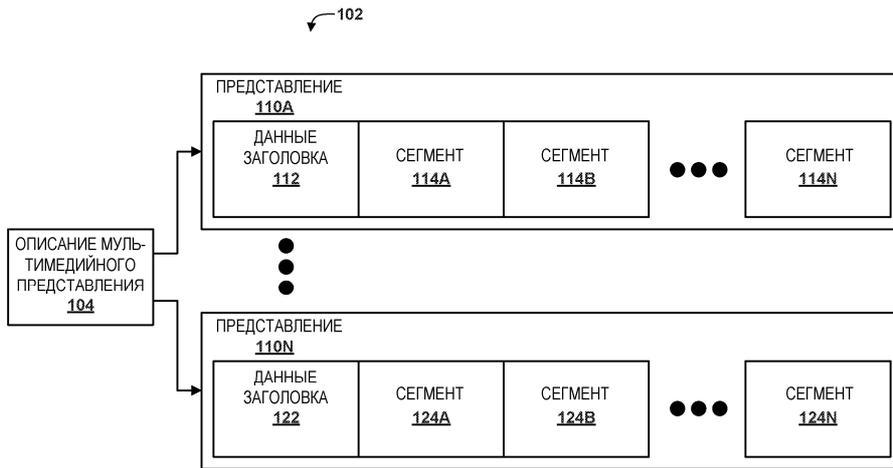
Фиг. 2



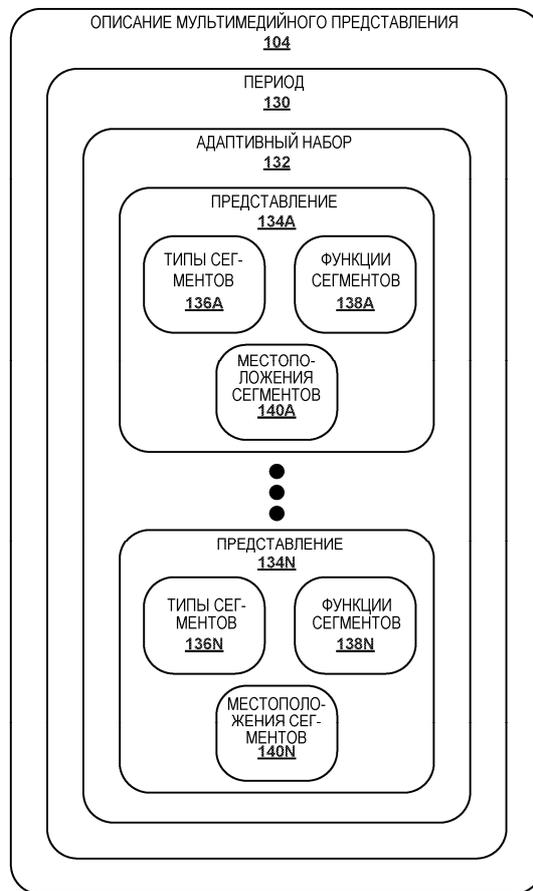
Фиг. 3



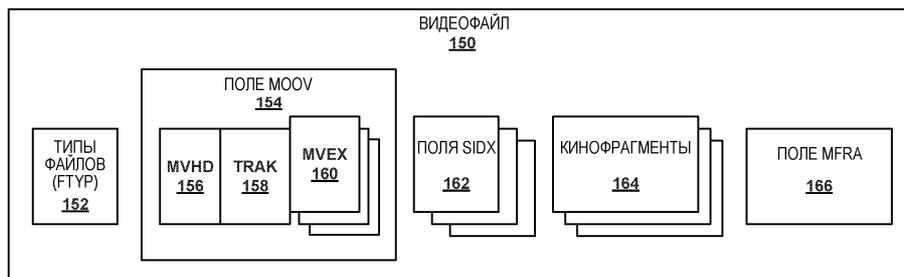
Фиг. 4



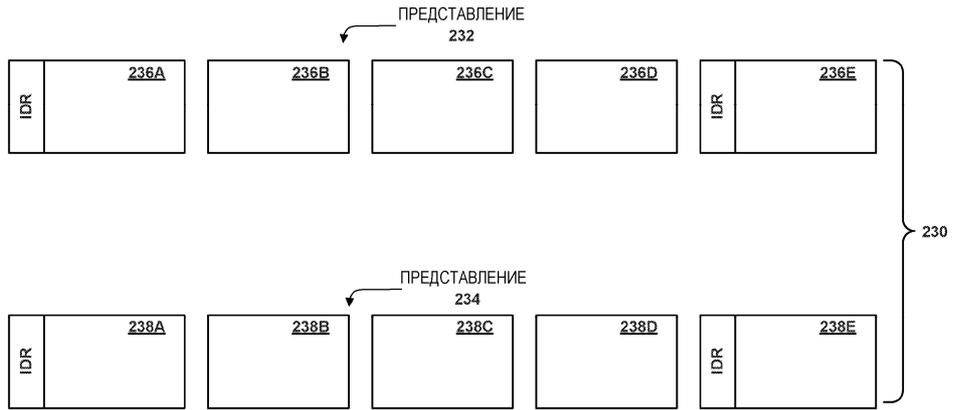
Фиг. 5A



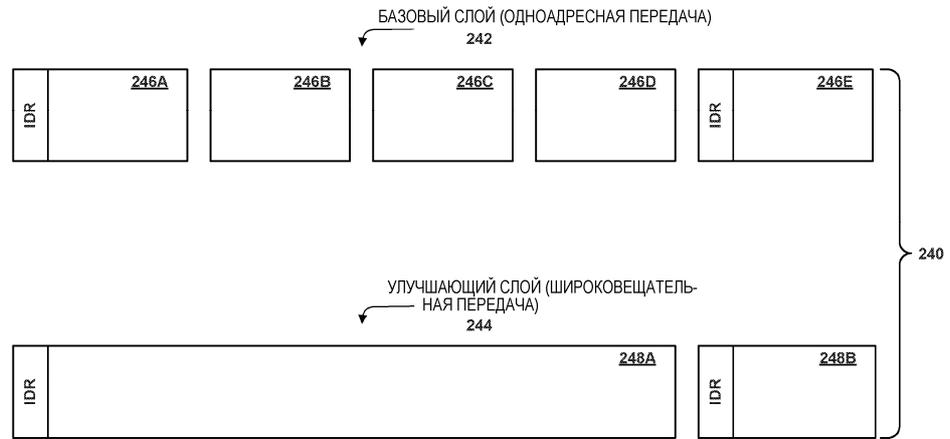
Фиг. 5B



Фиг. 6



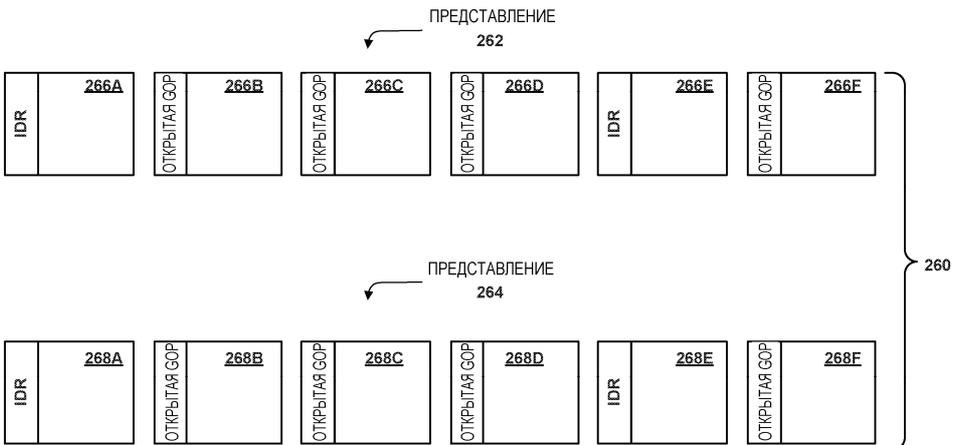
Фиг. 7



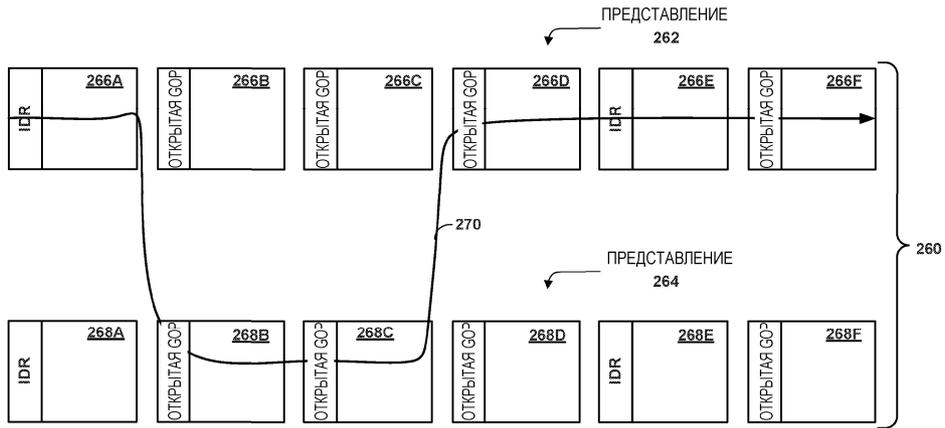
Фиг. 8



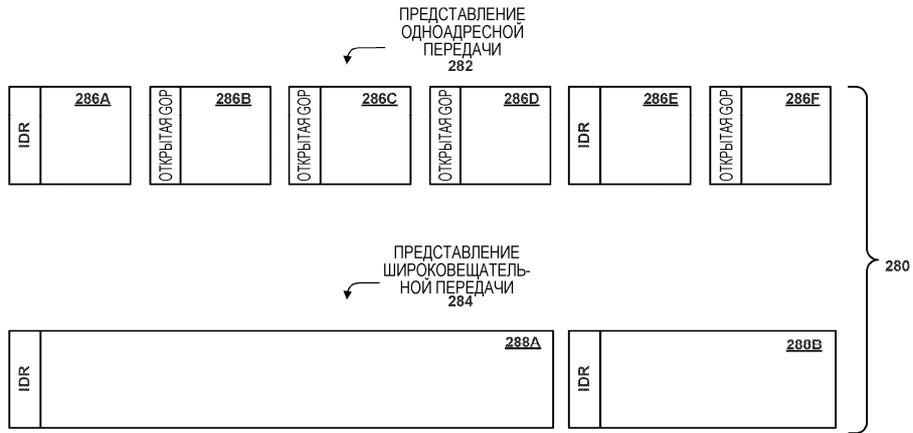
Фиг. 9



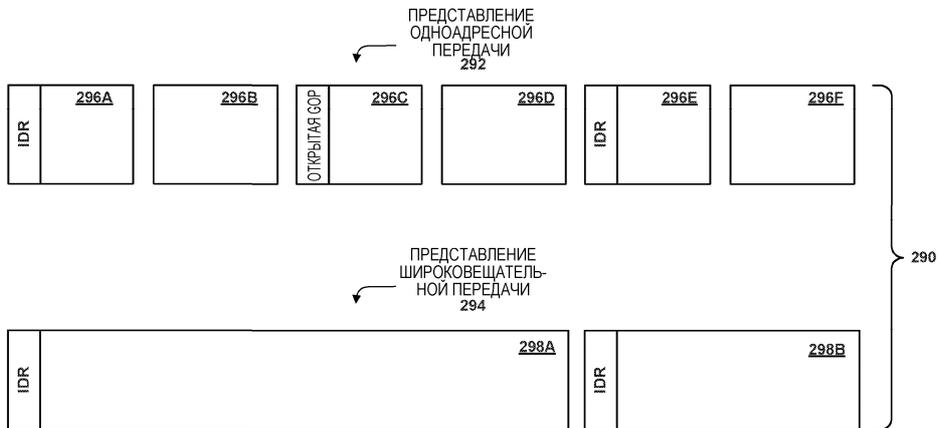
Фиг. 10



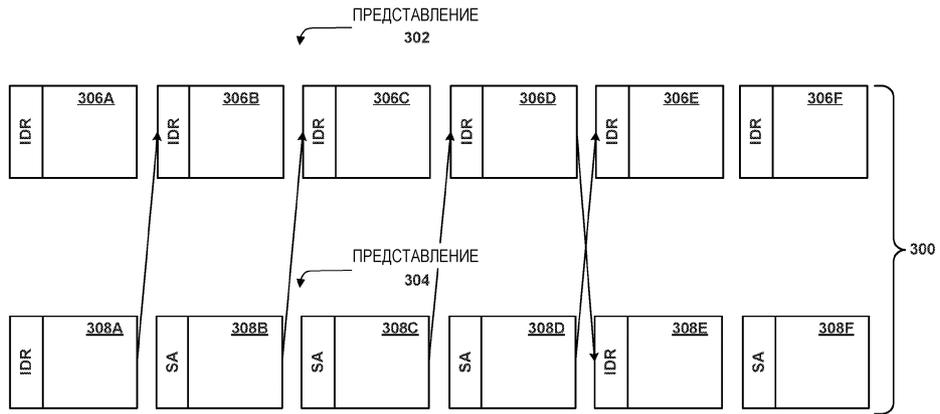
Фиг. 11



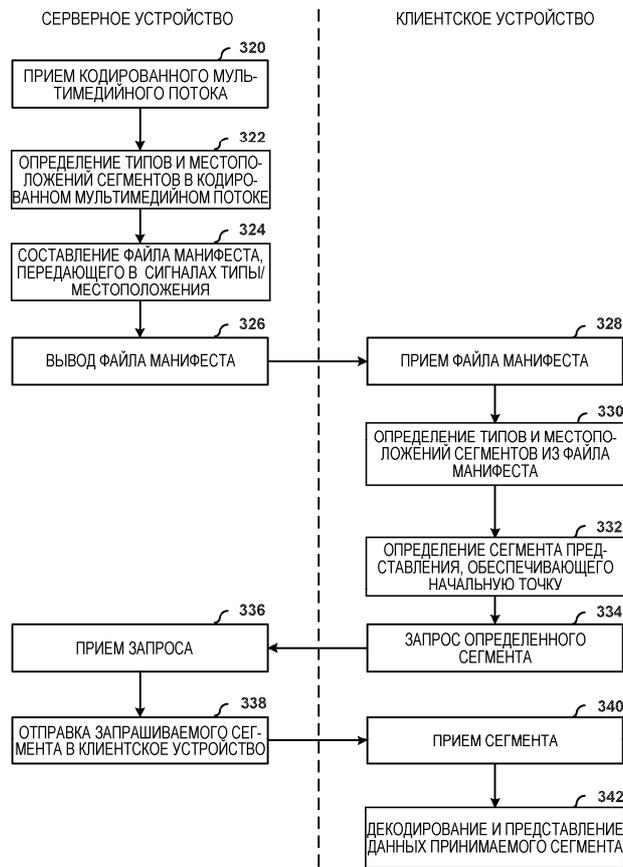
Фиг. 12



Фиг. 13



Фиг. 14



Фиг. 15